GENOMICS AND METABOLOMICS INSIGHTS INTO CARDIOVASCULAR DISEASE

Ruben Nathaniël Eppinga

Cover design by: Studio van der Velde

The figure on the cover is based on an error of one of the quality control plots (effect allele frequency) of a genome-wide associaton study, which by accident had the shape of a double helix.

Layout and printed by: Optima Grafische Communicatie (www.ogc.nl)

ISBN: 978-94-034-0529-2

Printing of this thesis was financially supported by: University of Groningen, Genzyme, Graduate School of Medical Sciences (GSMS), ChipSoft, Guerbet B.V, Sanofi.

rijksuniversiteit
groningen

GENOMICS AND METABOLOMICS INSIGHTS INTO CARDIOVASCULAR DISEASE

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 16 mei 2018 om 14.30 uur

door

Ruben Nathaniël Eppinga
geboren op 15 april 1984
te Smallingerland

Promotores
Prof. dr. P. van der Harst
Prof. dr. W.H. van Gilst

Copromotor
Dr. N. Verweij

Beoordelingscommissie
Prof. dr. P.I.W. de Bakker
Prof. dr. B.H.Ch. Stricker
Prof. dr. M.P. van den Berg

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

Paranimfen

Mevr. H.D. Eppinga

Mevr. H.T. Hartman

CONTENTS

Chapter 1	Introduction	9
PART I	**GENOMICS**	
Chapter 2	Identification of Genomic Loci Associated with Resting Heart Rate and Shared Genetic Predictors with All-Cause Mortality	21
Chapter 3	Telomere length and Risk of Cardiovascular Disease and Cancer	61
Chapter 4	Identification of 15 Novel Risk Loci for Coronary Artery Disease and Genetic Risk of Recurrent Events, Atrial Fibrillation and Heart Failure	67
PART II	**METABOLOMICS**	
Chapter 5	Effect of Metformin Treatment on Lipoprotein Subfractions in Non-Diabetic Patients with Acute Myocardial Infarction: A Glycometabolic Intervention as Adjunct to Primary Coronary Intervention in ST Elevation Myocardial Infarction (GIPS-III) Trial	97
Chapter 6	Effect of Metformin on Metabolites and Relation with Myocardial Infarct Size and Left Ventricular Ejection Fraction After Myocardial Infarction	121
Chapter 7	Statin-effects on Metabolic Profiles: Data from the Prevend It Trial	139
Chapter 8	Summary and Discussion	159
Chapter 9	Appendix	167

Chapter 1

Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide[1]. Major cardiovascular disorders include coronary artery disease (CAD), hypertension, cerebrovascular disease and peripheral arterial disease. In the past 50 years, considerable progress has been made in the definition, identification and modification of CVD risk factors and the development of appropriate medical and interventional treatments, such as percutaneous coronary interventions and use of β-blockers. All of these measures have resulted in a decline in cardiovascular mortality[2]. However, despite these efforts and subsequent progress, the mechanisms underlying the presentation and pathophysiology of CVD are still poorly understood. Identifying new biological and causal pathways may allow the development of new therapeutic strategies, risk stratification and increase our understanding of CVD pathophysiology. In this thesis, genomics (**Part I**) and metabolomics (**Part II**) were applied to gain biological and clinical insight into CVD traits. To this end, three overarching methodologies were used: a) Genome wide association studies to identify new genetic loci and genes to better understand the pathophysiology of CVD; b) Mendelian randomization approaches to identify potential causal pathways in disease; and c) interrogate the human metabolome to study the downstream products of gene transcription and identify new biomarkers.

PART I GENOMICS IN CVD

The first draft of the human genome (Human Genome Project[3]) was produced over 15 years ago, which led to a deeper understanding of genetic contributions of common variants to CVD. Prior to this, genes had been associated with CVD via Mendelian association, although these are relatively rare and constitute only a small portion of clinical CVD. Examples include: familial hypercholesterolemia, dilated and hypertrophic cardiomyopathy, long-QT syndrome, and aortic aneurysms[4]. The majority of CVD, however, are polygenic, with many heritable and environmental contributory factors[5]. Prior to the completion of the draft of the human genome, efforts to identify the genetic causes of polygenic CVD were largely unsuccessful.

The introduction of genome wide association studies, which test genetic variants across the genome for their association with a disease or trait, has allowed hundreds of loci for numerous CVD and traits to be identified. The aim of this thesis is to expand this field of research and provide new insights into the genomics of the cardiovascular system. We use genome wide association studies to identify new genetic variants and thereby further our knowledge of genes. Several bioinformatic methods were applied to the associated genetic variants to identify potential biological pathways and mechanisms. We also perform Mendelian randomization analyses, which uses the genetic

variants identified by genome wide associations studies as instrumental variables to estimate the causal effects of risk factors on disease development and mortality[6].

In the first part of this thesis, we study the genomics of heart rate (**Chapter 2**), telomere length (TL) (**Chapter 3)** and CAD (**Chapter 4**). We focus on heart rate and telomere length, as these are predictors of important clinical outcomes, such as overall mortality and CVD. The cause or consequence of these associations is still unknown and remains a topic of ongoing debate. **Chapter 4** examines one of the major CVD outcomes and the major cause of morbidity and mortality worldwide - coronary artery disease[7].

In **Chapter 2**, we further our knowledge of genetic influence on resting heart rate by performing a genome-wide association study. Resting heart rate in humans is a well-established predictor of overall mortality in the general population[8-13], as well as in patients with hypertension[14], CAD[15], and heart failure[16]. To date, the association of heart rate with life expectancy or risk does not provide sufficient evidence for a (shared) causal relationship. In some conditions (e.g. heart failure), reduction of heart rate has been shown to lead to event reduction[17]. This provides evidence that heart rate is not just a risk marker or reflection of comorbidities, but a modifiable, causal risk factor[16]. However, in patients with e.g. CAD and hypertension, β-adrenergic receptor-blocking agents were not associated with lower risk of cardiovascular events beyond their effect on blood pressure[18,19]. Moreover, in patients with atrial fibrillation (permanent), lenient and strict rate control are equally effective[20], and heart rate reduction with ivabradine did not improve outcomes in patients with CAD[21]. A mechanistic explanation linking higher resting heart rate with increased mortality remains enigmatic. In **Chapter 2** we further explore the relationship between resting heart rate with cardiovascular risk factors, comorbidities as well as fatal and non-fatal outcomes by using the identified genetic variants from our genome wide association study as instrumental variables in Mendelian randomization analysis. Similarly, in **Chapter 3** of this thesis we focus on the potential causal pathways of TL. It is known that short TL has been associated with an increased risk of mortality[22] and several CVD including CAD[23], abdominal aneurysms[24], atherosclerosis[23,25], heart failure[25] as well as cardiovascular risk factors such as smoking[25], increased body mass index[26], hypertension[25], and diabetes[25]. TL is also strongly related to both age and sex[27] although the cause or consequence of these associations in these cross-sectional studies is a topic of ongoing debate. In this chapter, we study the causal relationship between genetically determined TL with CVD and cancer risk by using genetic variants associated with TL obtained from previously published studies as instrumental variables in a Mendelian randomization analysis. In addition to studying two important risk factors (resting heart rate and TL) that have a strong links with CVD such as CAD, we performed a similar study with CAD itself in **Chapter 4**. CAD is driven by a complex interplay of multiple genetic and environmental factors that jointly give rise to a plethora of molecular interactions, resulting in a complex and heterogeneous phe-

notype. Genome wide association studies have identified about 57 loci associated with myocardial infarction and CAD[28]. These have identified targets of known and novel CAD medication such as LDLR and HMGCR (HMG-coA reducatase inhibitors, statins), PCSK9 (PCSK9 inhibitors) and IL6R (Tocilizumab)[29,30]. The aim of **Chapter 4** is to expand the genetic regions associated with CAD, facilitate the identification of additional therapeutic targets and gain insights into the causal relationships between other cardiovascular phenotypes.

PART II METABOLOMICS IN CVD

While genetic variants are key components of heritability, they are relatively 'static' components. For chronic diseases, including CVD, a closer examination of the disease process to identify biomarkers and understand mechanisms and pathophysiology of CVD may be valuable. ProBNP and troponin are important biomarkers of CVD and are used in the clinic for diagnosis and decision management. Metabolomics provides new opportunities to find novel biomarkers in CVD. Moreover, because metabolites represent downstream products of gene transcription, it also provides new opportunities to study biological pathways in disease. Metabolomics is a relatively novel field in 'omics' sciences, which uses high-throughput technologies, such as nuclear magnetic resonance (NMR) spectroscopy, to concurrently quantify a large number of small molecules in different tissues. Metabolic profiling has been successful in improving diagnosis and prediction of CV events[31,32] and differentiate heart failure patients from healthy controls[33]. Metabolic profiling may thus help identify novel biomarkers in CVD. To this end, we created metabolite profiles by measuring a large number of small molecules, including lipoprotein subfractions and lipid related measures, glycolysis related metabolites, amino-acids, ketone bodies, fluid balance related metabolites and an inflammatory marker using NMR spectroscopy in the GIPS-III and PREVEND-IT studies. In **Part II**, we study the changes involved in metabolites in patients with an acute myocardial infarction (AMI) and the effects of statin therapy on metabolite profiles using NMR spectroscopy.

By using NMR spectroscopy in the GIPS-III study, we tested the extent to which metabolomics - looking at lipoprotein subfractions (**Chapter 5**) and metabolic profiles (**Chapter 6**) - can predict left ventricular ejection fraction and infarct size after an AMI, both key predictors of long-term prognosis[34,35]. Furthermore, we investigated the effect of statin treatment - a critical treatment initiated shortly after AMI - on metabolic profiles and cardiovascular risk reduction (**Chapter 7**) by using data from PREVEND-IT: a randomized, double-blind, placebo-controlled study.

The genetic and metabolomic studies of Part I and II complement each other and may help us better understand the pathophysiology of cardiovascular disease, obtain new insights into risk prediction and provide new targets for therapy.

REFERENCES

1. WHO: Cardiovascular diseases (CVDs). http://www.who.int/mediacentre/factsheets/fs317/en/. Updated May, 2017.
2. Mensah GA, Wei GS, Sorlie PD, et al. Decline in cardiovascular mortality: Possible causes and implications. *Circ Res*. 2017;120(2):366-380.
3. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
4. Nabel EG. Cardiovascular disease. *N Engl J Med*. 2003;349(1):60-72.
5. Lee DS, Pencina MJ, Benjamin EJ, et al. Association of parental heart failure with risk of heart failure in offspring. *N Engl J Med*. 2006;355(2):138-147.
6. Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: Where are we now and where are we going? *Int J Epidemiol*. 2015;44(2):379-388.
7. Task Force Members, Montalescot G, Sechtem U, et al. 2013 ESC guidelines on the management of stable coronary artery disease: The task force on the management of stable coronary artery disease of the european society of cardiology. *Eur Heart J*. 2013;34(38):2949-3003.
8. Dyer AR, Persky V, Stamler J, et al. Heart rate as a prognostic factor for coronary heart disease and mortality: Findings in three chicago epidemiologic studies. *Am J Epidemiol*. 1980;112(6):736-749.
9. Kannel WB, Kannel C, Paffenbarger RS,Jr, Cupples LA. Heart rate and cardiovascular mortality: The framingham study. *Am Heart J*. 1987;113(6):1489-1494.
10. Gillum RF, Makuc DM, Feldman JJ. Pulse rate, coronary heart disease, and death: The NHANESI epidemiologic follow-up study. *Am Heart J*. 1991;121(1 Pt 1):172-177.
11. Greenland P, Daviglus ML, Dyer AR, et al. Resting heart rate is a risk factor for cardiovascular and noncardiovascular mortality: The chicago heart association detection project in industry. *Am J Epidemiol*. 1999;149(9):853-862.
12. Kristal-Boneh E, Silber H, Harari G, Froom P. The association of resting heart rate with cardiovascular, cancer and all-cause mortality. eight year follow-up of 3527 male israeli employees (the CORDIS study). *Eur Heart J*. 2000;21(2):116-124.
13. Reunanen A, Karjalainen J, Ristola P, Heliovaara M, Knekt P, Aromaa A. Heart rate and mortality. *J Intern Med*. 2000;247(2):231-239.
14. Kolloch R, Legler UF, Champion A, et al. Impact of resting heart rate on outcomes in hypertensive patients with coronary artery disease: Findings from the INternational VErapamil-SR/trandolapril STudy (INVEST). *Eur Heart J*. 2008;29(10):1327-1334.
15. Diaz A, Bourassa MG, Guertin MC, Tardif JC. Long-term prognostic value of resting heart rate in patients with suspected or proven coronary artery disease. *Eur Heart J*. 2005;26(10):967-974.
16. Bohm M, Swedberg K, Komajda M, et al. Heart rate as a risk factor in chronic heart failure (SHIFT): The association between heart rate and outcomes in a randomised placebo-controlled trial. *Lancet*. 2010;376(9744):886-894.
17. Swedberg K, Komajda M, Bohm M, et al. Ivabradine and outcomes in chronic heart failure (SHIFT): A randomised placebo-controlled study. *Lancet*. 2010;376(9744):875-885.
18. Bangalore S, Steg G, Deedwania P, et al. Beta-blocker use and clinical outcomes in stable outpatients with and without coronary artery disease. *JAMA*. 2012;308(13):1340-1349.
19. Messerli FH, Grossman E, Goldbourt U. Are beta-blockers efficacious as first-line therapy for hypertension in the elderly? A systematic review. *JAMA*. 1998;279(23):1903-1907.
20. Van Gelder IC, Groenveld HF, Crijns HJ, et al. Lenient versus strict rate control in patients with atrial fibrillation. *N Engl J Med*. 2010;362(15):1363-1373.

21. Fox K, Ford I, Steg PG, et al. Ivabradine in stable coronary artery disease without clinical heart failure. *N Engl J Med*. 2014;371(12):1091-1099.
22. Oeseburg H, de Boer RA, van Gilst WH, van der Harst P. Telomere biology in healthy aging and disease. *Pflugers Arch*. 2010;459(2):259-268.
23. Samani NJ, van der Harst P. Biological ageing and cardiovascular disease. *Heart*. 2008;94(5):537-539.
24. Atturu G, Brouilette S, Samani NJ, London NJ, Sayers RD, Bown MJ. Short leukocyte telomere length is associated with abdominal aortic aneurysm (AAA). *Eur J Vasc Endovasc Surg*. 2010;39(5):559-564.
25. Wong LS, de Boer RA, Samani NJ, van Veldhuisen DJ, van der Harst P. Telomere biology in heart failure. *Eur J Heart Fail*. 2008;10(11):1049-1056.
26. Nordfjäll K, Eliasson M, Stegmayr B, Melander O, Nilsson P, Roos G. Telomere length is associated with obesity parameters but with a gender difference. *Obesity*. 2008;16(12):2682-2689.
27. - Gardner M, - Bann D, - Wiley L, et al. - Gender and telomere length: Systematic review and meta-analysis. - *Experimental gerontology*. :- 15.
28. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121-1130.
29. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nat Rev Drug Discov*. 2013;12(8):581-594.
30. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium, Swerdlow DI, Holmes MV, et al. The interleukin-6 receptor as a target for prevention of coronary heart disease: A mendelian randomisation analysis. *Lancet*. 2012;379(9822):1214-1224.
31. Vaarhorst AA, Verhoeven A, Weller CM, et al. A metabolomic profile is associated with the risk of incident coronary heart disease. *Am Heart J*. 2014;168(1):45-52.e7.
32. Wurtz P, Raiko JR, Magnussen CG, et al. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J*. 2012;33(18):2307-2316.
33. Wang J, Li Z, Chen J, et al. Metabolomic identification of diagnostic plasma biomarkers in humans with chronic heart failure. *Mol Biosyst*. 2013;9(11):2618-2626.
34. El Aidi H, Adams A, Moons KG, et al. Cardiac magnetic resonance imaging findings and the risk of cardiovascular events in patients with recent myocardial infarction or suspected or known coronary artery disease: A systematic review of prognostic studies. *J Am Coll Cardiol*. 2014;63(11):1031-1045.
35. Wu E, Ortiz JT, Tejedor P, et al. Infarct size by contrast enhanced cardiac magnetic resonance is a stronger predictor of outcomes than left ventricular ejection fraction or end-systolic volume index: Prospective cohort study. *Heart*. 2008;94(6):730-736.

Part I

Genomics

Chapter 2

Identification of Genomic Loci Associated with Resting Heart Rate and Shared Genetic Predictors with All-Cause Mortality

Ruben N. Eppinga, Yanick Hagemeijer, Stephen Burgess, David A. Hinds, Kari Stefansson, Daniel F. Gudbjartsson, Dirk J. van Veldhuisen, Patricia B. Munroe, Niek Verweij, Pim van der Harst

Adapted from Nat Genet. 2016 Dec;48(12):1557-1563

ABSTRACT

Resting heart rate is a heritable trait correlated with lifespan. Little is known about the genetic contribution of resting heart rate and its relationship with mortality. We performed a genome-wide association discovery and replication analysis starting with 19.9 million genetic variants and studying up to 265,046 individuals to identify 64 loci associated with resting heart rate ($P<5\times10^{-8}$), 46 of these were novel. We then used the genetic variants identified to study the association between resting heart rate and all-cause mortality. We observed that a genetically predicted resting heart rate of 5 beats per minute was associated with a 20% increased mortality risk (hazard ratio 1.20, 95% CI of 1.11-1.28, $P=8.20\times10^{-7}$) translating to a 2.9 years reduction in life expectancy for males and 2.6 years for females. Our findings provide novel evidence for shared genetic predictors of resting heart rate and all-cause mortality.

INTRODUCTION

Among mammals, there exists an inverse semilogarithmic relation between resting heart rate and life expectancy with only the human species deviating from this line[1,2]. In humans, resting heart rate is a well-established predictor of overall mortality in the general population[3-8], as well as in patients with hypertension[9], coronary artery disease (CAD)[10], and heart failure[11]. The association of heart rate with life expectancy or risk does not provide sufficient evidence for a shared or causal relationship. Heart rate is regulated by complex interactions of biological systems, including the autonomous nervous and hormonal systems[12]. In addition, resting heart rate is associated with many other cardiovascular risk factors, including blood pressure, smoking, glucose metabolism, lipids, C-reactive protein, metabolic syndrome, body mass index, and diabetes mellitus[13-16]. In some conditions, including heart failure, reduction of heart rate has been directly demonstrated to lead to event reduction providing evidence that heart rate is indeed a modifiable causal risk factor and not just a risk marker or a reflection of comorbidities[11]. However, in patients with CAD and hypertension, β-adrenergic receptor-blocking agent (beta-blockers) were not associated with lower risk of cardiovascular events beyond its effect on blood pressure[17,18]; in patients with permanent atrial fibrillation, lenient rate control is as effective as strict rate control[19], and heart rate reduction with ivabradine did not improve outcomes in patients with CAD[20], though it does improve outcomes in patients with heart failure [21]. No mechanistic explanation linking higher resting heart rate with increased mortality has emerged. To further our knowledge on genes influencing resting heart rate we performed a genome-wide association study (GWAS) on 134,251 participants from UK Biobank[22] and replicated our findings in 130,795 additional individuals. Using the identified genetic variants as instrumental variables we explored the relationship between resting heart rate with cardiovascular risk factors, comorbidities and fatal and non-fatal outcomes. Bioinformatic analyses of associated variants were also undertaken to identify potential biological pathways and mechanisms.

We studied 134,251 individuals participating in UK Biobank. The average age was 56.6 years (interquartile range (IQR) 50 to 63), and 47.2% of the participants were male. Baseline characteristics are presented in Table 1 and Supplementary Table 1. The median duration of follow-up for mortality was 4.9 years (IQR 4.3 to 5.5 years) and there were 2,364 mortality events in total. Incidence rate 3.6 events (95% confidence interval (CI) 3.4 to 3.7 events) per 1000 person-years of follow-up.

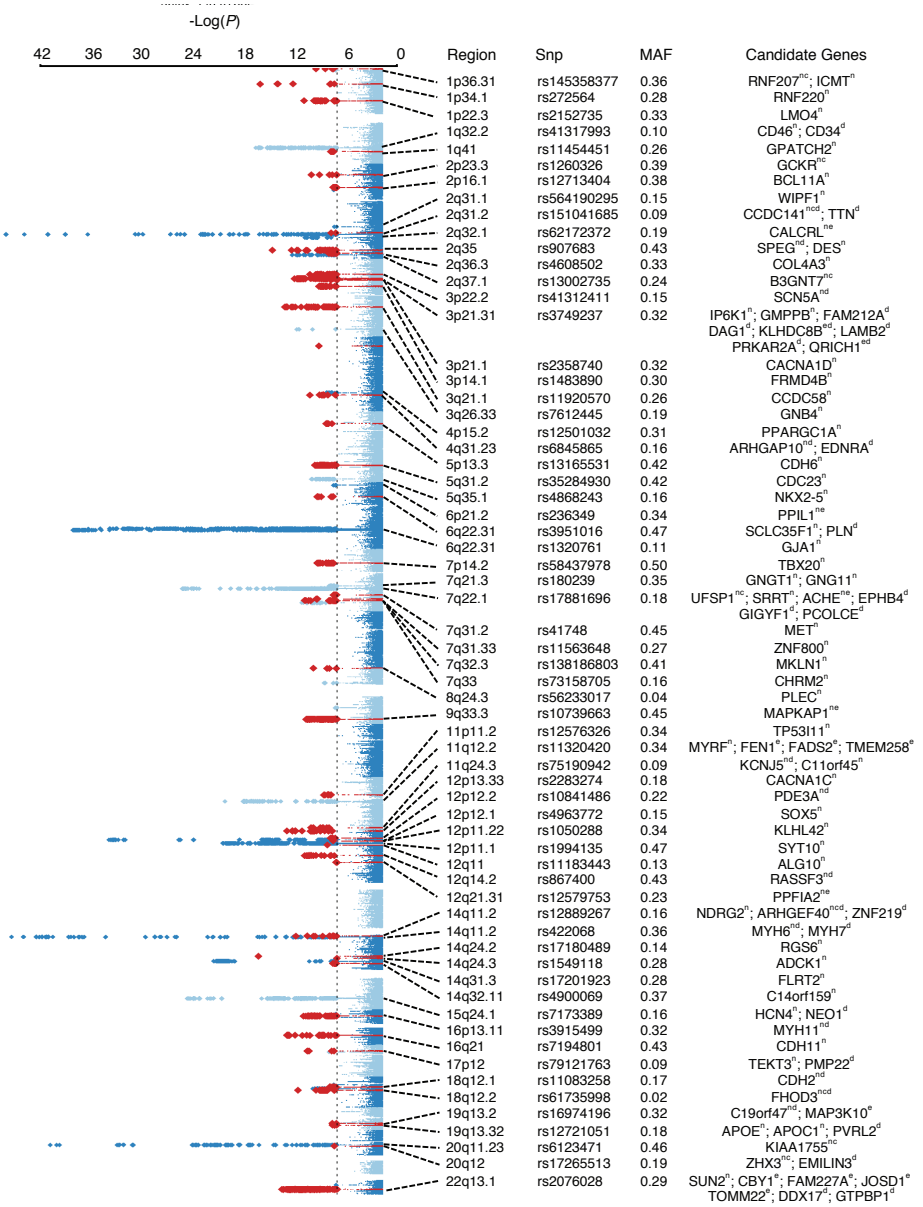
Table 1. Baseline characteristics of participants

	All (N=134,251)	SD or percentage (%)	Healthy Individuals (N=11,405)	SD or percentage (%)
Age	56.6	8.0	53.7	7.6
Sex (Male)	63,349	47.2%	5,993	52.5%
Body-mass index	27.5	4.8	26.3	4.0
Resting heart rate	69.5	11.1	68.3	10.5
Blood pressure				
Systolic	138.0	18.6	135.5	17.7
Diastolic	82.3	10.1	81.6	9.8
Ethnicity				
Asian	2,478	1.8%	248	2.2%
Black	1,734	1.3%	173	1.5%
Mixed	684	0.5%	52	0.5%
White	127,919	95.3%	10,797	94.7%
Other/ undefined	1,436	1.1%	135	1.2%
Smoking current	16,708	12.4%	1,390	8.3%
Medical History				
Hypertension	38,339	28.6%	0	0%
Diabetes	7,419	5.5%	0	0%
Myocardial Infarction	3,395	2.5%	0	0%
Heart failure	720	0.5%	0	0%
Atrial fibrillation / flutter	2,048	1.5%	0	0%
Supraventricular tachycardia	425	0.3%	0	0%
Device implantation	399	0.3%	0	0%
Medication			0	0%
Beta-blockers	9,526	7.8%	0	0%
Calcium-channel blockers	9,797	8.0%	0	0%

Abbreviations; SD, Standard Deviation.

RESULTS

In UK Biobank we identified genetic variants at 76 loci associated with resting heart rate at $P<5\times10^{-8}$ (Figure 1, Table 2, Supplementary Table 2, Supplementary Figures 1-3). 64 of these loci replicated in 130,795 individuals derived from 4 cohorts , and 46 loci have not been previously reported as associated with resting heart rate[23]. The genetic variants at the 64 loci were well imputed with an info >0.9, except one (rs11183443) which had an information measure of 0.30. At 11 loci we found evidence for multiple independent associations with resting heart rate in conditional analyses (Supplementary Table 3). As



Region	Snp	MAF	Candidate Genes
1p36.31	rs145358377	0.36	RNF207[nc]; ICMT[n]
1p34.1	rs272564	0.28	RNF220[n]
1p22.3	rs2152735	0.33	LMO4[n]
1q32.2	rs41317993	0.10	CD46[n]; CD34[d]
1q41	rs11454451	0.26	GPATCH2[n]
2p23.3	rs1260326	0.39	GCKR[nc]
2p16.1	rs12713404	0.38	BCL11A[n]
2q31.1	rs564190295	0.15	WIPF1[n]
2q31.2	rs151041685	0.09	CCDC141[ncd]; TTN[d]
2q32.1	rs62172372	0.19	CALCRL[ne]
2q35	rs907683	0.43	SPEG[nd]; DES[n]
2q36.3	rs4608502	0.33	COL4A3[n]
2q37.1	rs13002735	0.24	B3GNT7[nc]
3p22.2	rs41312411	0.15	SCN5A[nd]
3p21.31	rs3749237	0.32	IP6K1[n]; GMPPB[n]; FAM212A[d] DAG1[d]; KLHDC8B[nd]; LAMB2[d] PRKAR2A[e]; QRICH1[ed]
3p21.1	rs2358740	0.32	CACNA1D[n]
3p14.1	rs1483890	0.30	FRMD4B[n]
3q21.1	rs11920570	0.26	CCDC58[n]
3q26.33	rs7612445	0.19	GNB4[n]
4p15.2	rs12501032	0.31	PPARGC1A[n]
4q31.23	rs6845865	0.16	ARHGAP10[nd]; EDNRA[d]
5p13.3	rs13165531	0.42	CDH6[n]
5q31.2	rs35284930	0.42	CDC23[n]
5q35.1	rs4868243	0.16	NKX2-5[n]
6p21.2	rs236349	0.34	PPIL1[ne]
6q22.31	rs3951016	0.47	SCLC35F1[n]; PLN[d]
6q22.31	rs1320761	0.11	GJA1[n]
7p14.2	rs58437978	0.50	TBX20[n]
7q21.3	rs180239	0.35	GNGT1[n]; GNG11[n]
7q22.1	rs17881696	0.18	UFSP1[nc]; SRRT[n]; ACHE[ne]; EPHB4[d] GIGYF1[d]; PCOLCE[d]
7q31.2	rs41748	0.45	MET[n]
7q31.33	rs11563648	0.27	ZNF800[n]
7q32.3	rs138186803	0.41	MKLN1[n]
7q33	rs73158705	0.16	CHRM2[n]
8q24.3	rs56233017	0.04	PLEC[n]
9q33.3	rs10739663	0.45	MAPKAP1[ne]
11p11.2	rs12576326	0.34	TP53I11[n]
11q12.2	rs11320420	0.34	MYRF[n]; FEN1[e]; FADS2[e]; TMEM258[e]
11q24.3	rs75190942	0.09	KCNJ5[nd]; C11orf45[n]
12p13.33	rs2283274	0.18	CACNA1C[n]
12p12.2	rs10841486	0.22	PDE3A[nd]
12p12.1	rs4963772	0.15	SOX5[n]
12p11.22	rs1050288	0.34	KLHL42[n]
12p11.1	rs1994135	0.47	SYT10[n]
12q11	rs11183443	0.13	ALG10[n]
12q14.2	rs867400	0.43	RASSF3[nd]
12q21.31	rs12579753	0.23	PPFIA2[ne]
14q11.2	rs12889267	0.16	NDRG2[n]; ARHGEF40[ncd]; ZNF219[d]
14q11.2	rs422068	0.36	MYH6[nd]; MYH7[d]
14q24.2	rs17180489	0.14	RGS6[n]
14q24.3	rs1549118	0.28	ADCK1[n]
14q31.3	rs17201923	0.28	FLRT2[n]
14q32.11	rs4900069	0.37	C14orf159[n]
15q24.1	rs7173389	0.16	HCN4[n]; NEO1[d]
16p13.11	rs3915499	0.32	MYH11[nd]
16q21	rs7194801	0.43	CDH11[n]
17p12	rs79121763	0.09	TEKT3[n]; PMP22[d]
18q12.1	rs11083258	0.17	CDH2[nd]
18q12.2	rs61735998	0.02	FHOD3[ncd]
19q13.2	rs16974196	0.32	C19orf47[nd]; MAP3K10[e]
19q13.32	rs12721051	0.18	APOE[n]; APOC1[n]; PVRL2[d]
20q11.23	rs6123471	0.46	KIAA1755[nc]
20q12	rs17265513	0.19	ZHX3[nc]; EMILIN3[d]
22q13.1	rs2076028	0.29	SUN2[n]; CBY1[e]; FAM227A[e]; JOSD1[e] TOMM22[e]; DDX17[d]; GTPBP1[d]

Figure 1. Genomewide $-\log_{10}(P)$ plot and effects for significant loci

Genomewide $-\log_{10}(P)$ plots are shown for heart rate. Blue indicates previously identified genetic variants within loci reaching genome-wide significance; red indicates novel genetic variants within loci reaching genome-wide significance (±1 Mb of lowest P value). The dashed line indicates the genomewide significance threshold ($P=5\times10^{-8}$). Candidate genes are listed along with strategies used to identify them: n, nearest; c, coding; nonsynonymous variant; e, eQTL; d, DEPICT tool.

Table 2. Results of the newly identified loci that showed association with heart rate at genome-wide significance ($P<5\times10^{-8}$)

			Alleles		Discovery (UK Biobank)				Replication			Meta-analysis				
Variant	Chr.	Pos.[a]	Non-coded	Coded	EAF	β	SE	*P*	β	SE	*P*	β	SE	*P*	N	Candidate Genes
rs145358377	1	6272136	G	GA	0.36	-0.29	0.04	1.69×10^{-10}	-0.182	0.077	1.78×10^{-2}	-0.259	0.039	1.94×10^{-11}	265,046	RNF207[nc]; ICMT[n]
rs272564	1	45012273	A	C	0.28	0.41	0.05	5.02×10^{-17}	0.271	0.058	3.30×10^{-6}	0.351	0.037	4.51×10^{-21}	265,046	RNF220[n]
rs2152735	1	87893132	G	A	0.33	-0.32	0.05	6.48×10^{-12}	-0.291	0.056	2.06×10^{-7}	-0.306	0.036	7.23×10^{-18}	265,046	LMO4[n]
rs11454451	1	217722890	C	CT	0.26	0.28	0.05	8.81×10^{-9}	0.216	0.059	2.45×10^{-4}	0.256	0.038	1.29×10^{-11}	265,046	GPATCH2[n]
rs1260326	2	27730940	T	C	0.39	-0.29	0.04	4.54×10^{-11}	-0.256	0.054	1.75×10^{-6}	-0.275	0.034	4.29×10^{-16}	265,046	GCKR[nc]
rs12713404	2	60006705	G	T	0.38	-0.26	0.05	1.75×10^{-8}	-0.120	0.053	2.42×10^{-2}	-0.199	0.035	9.33×10^{-9}	265,046	BCL11A[n]
rs564190295	2	175547672	G	GCCGCC GCCCCC	0.15	-0.36	0.06	1.00×10^{-8}	-0.344	0.142	1.57×10^{-2}	-0.355	0.057	4.95×10^{-10}	197,184	WIPF1[n]
rs907683	2	220299541	G	T	0.43	-0.35	0.04	1.27×10^{-15}	-0.296	0.061	1.10×10^{-6}	-0.334	0.036	1.02×10^{-20}	265,046	SPEG[nd]; DES[n]
rs4608502	2	228134155	T	C	0.33	0.27	0.05	5.44×10^{-9}	0.221	0.055	6.02×10^{-5}	0.249	0.035	1.85×10^{-12}	265,046	COL4A3[n]
rs41312411	3	38621237	C	G	0.15	-0.40	0.06	3.31×10^{-11}	-0.192	0.076	1.13×10^{-2}	-0.320	0.047	1.34×10^{-11}	265,046	SCN5A[b,d]
rs3749237	3	49770032	G	A	0.32	0.33	0.05	5.18×10^{-13}	0.150	0.056	7.40×10^{-3}	0.258	0.035	3.09×10^{-13}	265,046	IP6K1[n]; GMPPB[n]; FAM212A[d]; DAG1[d];KLHDC8B[ed]; LAMB2[d]; PRKAR2A[d]; QRICH1[ed]
rs2358740	3	53455569	G	T	0.32	-0.26	0.05	9.24×10^{-9}	-0.128	0.055	2.03×10^{-2}	-0.208	0.035	3.58×10^{-9}	265,046	CACNA1D[n]
rs1483890	3	69410725	A	G	0.30	0.29	0.05	3.56×10^{-10}	0.272	0.056	1.38×10^{-6}	0.284	0.036	2.54×10^{-15}	265,046	FRMD4B[n]
rs11920570	3	122090102	G	A	0.26	0.37	0.05	3.91×10^{-14}	0.127	0.058	2.75×10^{-2}	0.268	0.037	5.18×10^{-13}	265,046	CCDC58[n]
rs12501032	4	23951018	C	G	0.31	0.29	0.05	3.65×10^{-10}	0.278	0.057	9.80×10^{-7}	0.288	0.036	1.83×10^{-15}	265,046	PPARGC1A[n]
rs6845865	4	148974602	T	C	0.16	-0.38	0.06	3.16×10^{-11}	-0.281	0.072	9.07×10^{-5}	-0.342	0.045	2.25×10^{-14}	265,046	ARHGAP10[nd]; EDNRA[d]
rs13165531	5	30888583	A	T	0.42	-0.26	0.04	2.75×10^{-9}	-0.166	0.053	1.65×10^{-3}	-0.221	0.034	4.31×10^{-11}	265,046	CDH6[n]

Table 2. Results of the newly identified loci that showed association with heart rate at genome-wide significance ($P<5\times10^{-8}$) (continued)

			Alleles			Discovery (UK Biobank)			Replication			Meta-analysis				
Variant	Chr.	Pos.[a]	Non-coded	Coded	EAF	β	SE	*P*	β	SE	*P*	β	SE	*P*	N	Candidate Genes
rs1468333[#]	5	137552970	T	C	0.63	-0.27	0.04	1.23×10^{-9}	-0.233	0.054	1.52×10^{-5}	-0.255	0.034	9.53×10^{-14}	265,046	CDC23[n]
rs236349	6	36820565	A	G	0.34	0.29	0.05	2.46×10^{-10}	0.273	0.055	8.11×10^{-7}	0.281	0.035	1.01×10^{-15}	265,046	PPIL1[ne]
rs58437978	7	35258277	T	C	0.50	-0.27	0.04	2.26×10^{-10}	-0.183	0.057	1.32×10^{-3}	-0.240	0.034	2.61×10^{-12}	265,046	TBX20[n]
rs41748	7	116446573	T	G	0.45	-0.24	0.04	1.90×10^{-8}	-0.120	0.052	2.23×10^{-2}	-0.193	0.033	7.14×10^{-9}	265,046	MET[n]
rs11563648	7	126970046	G	C	0.27	-0.31	0.05	1.79×10^{-10}	-0.121	0.058	3.74×10^{-2}	-0.231	0.037	4.42×10^{-10}	265,046	ZNF800[n]
rs138186803	7	130965408	AT	A	0.41	-0.30	0.04	7.81×10^{-12}	-0.550	0.107	2.70×10^{-7}	-0.333	0.040	1.27×10^{-16}	197,184	MKLN1[n]
rs56233017	8	144981488	G	A	0.04	-0.68	0.11	8.41×10^{-11}	-0.637	0.135	2.49×10^{-6}	-0.666	0.083	1.09×10^{-15}	265,046	PLEC[n]
rs10739663	9	128278739	A	G	0.45	-0.29	0.04	1.20×10^{-11}	-0.229	0.052	1.05×10^{-5}	-0.266	0.033	9.62×10^{-16}	265,046	MAPKAP1[ne]
rs12576326	11	44980383	A	G	0.34	0.27	0.05	1.40×10^{-9}	0.219	0.058	1.57×10^{-4}	0.253	0.036	1.20×10^{-12}	265,046	TP53I11[n]
rs75190942	11	128764571	C	A	0.09	-0.50	0.08	4.72×10^{-11}	-0.498	0.099	4.90×10^{-7}	-0.496	0.060	1.19×10^{-16}	265,046	KCNJ5[nd]; C11orf45[n]
rs2283274	12	2184466	G	C	0.18	-0.43	0.06	6.53×10^{-14}	-0.371	0.071	1.58×10^{-7}	-0.405	0.044	7.21×10^{-20}	265,046	CACNA1C[n]
rs10841486	12	20472202	T	C	0.22	-0.30	0.05	8.65×10^{-9}	-0.148	0.063	1.89×10^{-2}	-0.238	0.040	2.98×10^{-9}	265,046	PDE3A[nd]
rs1050288	12	27955296	C	T	0.34	-0.26	0.05	1.70×10^{-8}	-0.142	0.057	1.36×10^{-2}	-0.213	0.036	2.74×10^{-9}	265,046	KLHL42[n]
rs10880689[$]	12	37930102	A	G	0.60	0.20	0.04	4.65×10^{-6}	0.221	0.054	3.91×10^{-5}	0.208	0.034	8.10×10^{-10}	265,046	ALG10B[n]
rs867400	12	64976850	T	C	0.43	0.30	0.04	7.80×10^{-12}	0.301	0.053	1.05×10^{-8}	0.298	0.033	4.58×10^{-19}	265,046	RASSF3[nd]
rs12579753	12	82219376	C	T	0.23	-0.28	0.05	3.92×10^{-8}	-0.193	0.062	1.74×10^{-3}	-0.246	0.039	4.81×10^{-10}	265,046	PPFIA2[ne]
rs12889267	14	21542766	A	G	0.16	0.41	0.06	7.78×10^{-13}	0.421	0.073	7.89×10^{-9}	0.416	0.045	3.61×10^{-20}	265,046	NDRG2[n]; ARHGEF40[ncd]; ZNF219[d]
rs17180489	14	72885471	G	C	0.14	-0.52	0.06	3.14×10^{-17}	-0.370	0.132	5.01×10^{-3}	-0.490	0.055	9.15×10^{-19}	214,007	RGS6[n]
rs1549118	14	78379684	C	T	0.28	0.26	0.05	4.59×10^{-8}	0.113	0.057	4.80×10^{-2}	0.200	0.037	4.67×10^{-8}	265,046	ADCK1[n]
rs4900069	14	91583373	A	C	0.37	0.25	0.04	1.55×10^{-8}	0.125	0.054	2.14×10^{-2}	0.200	0.034	5.38×10^{-9}	265,046	C14orf159[n]

Table 2. Results of the newly identified loci that showed association with heart rate at genome-wide significance ($P<5\times10^{-8}$) (continued)

			Alleles			Discovery (UK Biobank)			Replication			Meta-analysis				
Variant	Chr.	Pos.[a]	Non-coded	Coded	EAF	β	SE	*P*	β	SE	*P*	β	SE	*P*	N	Candidate Genes
rs3915499	16	15910743	G	A	0.32	0.32	0.05	5.94×10^{-12}	0.284	0.056	3.72×10^{-7}	0.303	0.035	1.24×10^{-17}	265,046	MYH11[nd]
rs7194801	16	65286870	T	C	0.43	-0.33	0.04	6.78×10^{-14}	-0.240	0.052	4.49×10^{-6}	-0.291	0.033	3.58×10^{-18}	265,046	CDH11[n]
rs79121763	17	15195279	C	T	0.09	-0.52	0.08	1.53×10^{-11}	-0.376	0.110	6.59×10^{-4}	-0.471	0.063	7.17×10^{-14}	265,046	TEKT3[n]; PMP22[d]
rs11083258	18	25766218	A	C	0.17	-0.33	0.06	7.25×10^{-9}	-0.192	0.071	6.96×10^{-3}	-0.276	0.045	5.51×10^{-10}	265,046	CDH2[nd]
rs61735998	18	34289285	G	T	0.02	-0.98	0.14	1.39×10^{-12}	-0.593	0.176	7.74×10^{-4}	-0.834	0.109	2.06×10^{-14}	265,046	FHOD3[ncd]
rs16974196	19	40833470	G	A	0.32	0.26	0.05	1.36×10^{-8}	0.217	0.057	1.55×10^{-4}	0.244	0.036	1.11×10^{-11}	265,046	C19orf47[nd]; MAP3K10[e]
rs12721051	19	45422160	C	G	0.18	-0.32	0.06	1.40×10^{-8}	-0.241	0.071	6.45×10^{-4}	-0.287	0.044	5.23×10^{-11}	265,046	APOE[n]; APOC1[n]; PVRL2[d]
rs17265513	20	39832628	T	C	0.19	0.30	0.05	2.36×10^{-8}	0.146	0.066	2.78×10^{-2}	0.240	0.042	1.12×10^{-8}	265,046	ZHX3[nc]; EMILIN3[d]
rs2076028	22	39150450	G	A	0.29	-0.36	0.05	1.81×10^{-14}	-0.197	0.057	5.49×10^{-4}	-0.295	0.036	5.45×10^{-16}	265,046	SUN2[n]; CBY1[e]; FAM227A[e]; JOSD1[e]; TOMM22[e]; DDX17[d]; GTPBP1[d]

EAF, effect allele frequency; chr., chromosome; pos, position; SE, standard error. [a]Positions are according to 1000 Genomes phase 3, and allele coding is based on the positive strand. Candidate genes have been identified by one or multiple strategies: [b]nearest; [c]coding, nonsynonymous variant; [d]DEPICT tool; [e]eQTL. [f]Proxy of rs35284930, R^2=0.85. [g]Proxy of rs11183443, R^2=0.92.

expected, the magnitudes of the associations were small and ranged from 0.2 to 1.1 bpm per effect allele. Collectively, the total variance explained by the 64 loci for resting heart rate was 2.5%.

We studied the potential modifying effect of gender, beta-blockers and calcium-channel blockers on the association of genetic variants on resting heart rate but did not observe any significant interactions (Supplementary Table 4).

We summed the number of resting heart rate increasing alleles weighted for the strength of the association in the replication dataset to create a weighted genetic risk score (GRS) for each individual, and evaluated associations with cardiovascular measures. Genetically determined higher resting heart rate was associated with higher body-mass index and systolic and diastolic blood pressure, higher odds of having hypertension, active smoking behavior, experiencing supraventricular tachycardias and lower odds of device implantation (all $P<0.05$; Table 3). Shared heritability estimates are presented in Supplementary Table 5 and indicate correlations of resting heart rate with body-mass index, blood pressure, hypertension, diabetes, active smoking behavior, and myocardial infarction.

In a random-effects meta-analysis of the genetic variant-specific β_3 (the putative association between resting heart rate and outcome mediated through that variant) of all hypothesis generating loci ($P<1\times10^{-5}$) we observed a significant association between genetic variants associated with resting heart rate and all-cause mortality, translating to a relative increase of 20% in all-cause mortality risk per 5 bpm increase of resting heart rate (estimated hazard ratio (HR)=1.20, CI=1.11-1.28, $P=8.20\times10^{-7}$) (Table 4 and Supplementary Figure 4). When the number of genetic variants was restricted stepwise from $P<1\times10^{-5}$ to $P<5\times10^{-8}$, the HR decreased but remained significant (Table 4).

Next we calculated weighted and unweighted GRS and found similar associations with all-cause mortality (Table 4). Kaplan-Meier failure curves for all-cause mortality are shown in Supplementary Figure 5. There was no specific cause of death driving the association (Supplementary Table 6). We extrapolated a relative risk of 1.20 to life expectancy using the National Life Tables of the United Kingdom (Methods) and estimated a reduction of 2.9 years for males and 2.6 years for females per 5 bpm increase in resting heart rate.

A conceptual figure of the potential explanations of the observed association between genetic variants of heart rate and outcome is provided as Supplementary Figure 6. We performed several analyses to test for pleiotropic effects, identify confounders and mediators. First, we ruled out the possibility that extreme associations drive the genetic association with all-cause mortality by repeating the meta-analysis without the 12 genetic variants that each showed an association with mortality at $P<0.05$ (Table 4). Second, we adjusted for resting heart rate in the Cox regression model predicting all-cause mortality. The association of the genetic variants with all-cause mortality was abolished suggest-

ing the genetic association is mediated via resting heart rate (Table 4). Next, we adjusted for covariates observed to be associated with identified genetic variants in UK Biobank (Table 4). Introducing baseline body-mass index, diastolic blood pressure, hypertension, diabetes, active smoking, history of heart failure, supraventricular tachycardias, myocardial infarction, device implantation, beta-blockers and calcium channel-blockers did not affect the association between the genetic variants for heart rate and all-cause mortality (Table 4). Further, when we excluded all genetic variants that individually showed nominal significant association in UK Biobank ($P<0.05$) with any of the variables in Table 3, the association between the genetic variants for heart rate and all-cause mortality remained significant. Next, we considered potential confounders of variables not available in the UK Biobank cohort and performed multivariable Mendelian randomization (MR) to adjust for blood lipid levels (low-density lipoprotein, high-density lipoprotein, total cholesterol, triglycerides) and red blood cell (red blood cell count, packed cell volume mean corpuscular volume and hemoglobin count) variables. The adjustments did not attenuate the association of the heart rate-associated genetic variants with all-cause mortality (Table 4). The results of the MR-Egger method confirmed the absence of

Table 3. Association between genetically determined heart rate and cardiovascular profile using a weighted GRS

	Participants (N=134,251)	Percentage (%)	Estimated Association*	95% CI	*P* value
Body-mass index	134,251	100.0	0.14	0.08 to 0.20	2.24×10^{-6}
Blood pressure					
Systolic	134,217	99.0	-0.51	-0.30 to -0.72	2.55×10^{-6}
Diastolic	134,217	99.0	0.78	0.66 to 0.90	1.32×10^{-36}
Hypertension	39,996	29.8	1.04	1.01 to 1.07	4.41×10^{-3}
Diabetes	7,857	5.9	1.04	0.99 to 1.09	0.16
Smoking current	16,708	12.4	1.07	1.03 to 1.11	2.98×10^{-4}
Myocardial Infarction	3,848	2.9	0.99	0.92 to 1.07	0.80
Heart failure	1,131	0.8	1.14	0.99 to 1.31	0.06
Atrial fibrillation / flutter	2,780	2.1	1.01	0.93 to 1.10	0.79
Supraventricular tachycardia	546	0.4	1.28	1.05 to 1.56	0.02
Device implantation	482	0.4	0.80	0.66 to 0.96	0.02
Medication					
Beta-blockers	9,526	7.8	1.04	0.99 to 1.09	0.10
Calcium-channel blockers	9,797	8.0	1.02	0.98 to 1.07	0.34

* The effect estimates with 95% Confidence Interval (CI) estimated using weighted GRS (per 5 bpm increase in resting heart rate) are shown as odds ratios for categorical variables (hypertension, diabetes, smoking current, myocardial infarction, heart failure, atrial fibrillation / flutter, supraventricular tachycardia, device implantation, beta-blockers and calcium-channel blockers) and β estimates for quantitative variables (body-mass index, systolic and diastolic blood pressure).

Table 4. Association between genetically determined resting heart rate and all-cause mortality

Association with mortality	Number of GVs	Estimated Association HR*	95% CI	*P* value
Standard MR with all				
GVs ($P<10^{-2}$)	1980	1.19	1.14 to 1.23	3.77×10^{-19}
GVs ($P<10^{-3}$)	1739	1.19	1.14 to 1.23	5.91×10^{-19}
GVs ($P<10^{-4}$)	848	1.19	1.13 to 1.24	1.13×10^{-11}
GVs ($P<10^{-5}$)	272	1.20	1.11 to 1.28	8.20×10^{-7}
GVs ($P<10^{-6}$)	121	1.14	1.05 to 1.25	3.33×10^{-3}
GVs ($P<10^{-7}$)	82	1.13	1.02 to 1.25	1.46×10^{-2}
GVs ($P<5\times10^{-8}$)	76	1.11	1.00 to 1.22	5.01×10^{-2}
GVs ($P<10^{-5}$) excluding those associated ($P<0.05$) mortality	260	1.15	1.07 to 1.24	1.53×10^{-4}
GVs ($P<10^{-5}$) with adj. for resting heart rate	272	1.02	0.95 to 1.09	0.65
GVs ($P<10^{-5}$) with adj. for covariates#	272	1.18	1.10 to 1.27	4.69×10^{-6}
GVs ($P<10^{-5}$) excluding those associated ($P<0.05$) with variable#	55	1.29	1.09 to 1.53	3.66×10^{-3}
GVs ($P<10^{-5}$) betas estimated on 11,405 healthy individuals	272	1.14	1.07 to 1.23	6.85×10^{-5}
GVs ($P<10^{-5}$) betas estimated on 130,795 individuals from replication	269	1.11	1.01 to 1.22	2.70×10^{-2}
GRS weighted GVs ($P<10^{-5}$)	272	1.18	1.10 to 1.26	3.22×10^{-6}
GRS unweighted GVs ($P<10^{-5}$)	272	1.05	1.03 to 1.08	4.37×10^{-5}
Multivariable MR with adj. for covariates#	272	1.26	1.13 to 1.42	8.03×10^{-5}
Multivariable MR with adj. for lipid covariates$	209	1.18	1.09 to 1.27	1.99×10^{-5}
Multivariable MR with adj. for red blood cell covariates@	173	1.18	1.09 to 1.28	4.53×10^{-5}
MR-Egger method ($P<10^{-5}$)	272	1.21	1.05 to 1.40	8.00×10^{-3}

*Hazard ratio (HR) with 95% Confidence Interval (CI) estimated with standard Mendelian Randomization (MR) and weighted Genetic Risk Score (GRS) per 5 bpm and for unweighted GRS per 5 summed risk alleles; Genetic Variants (GVs); Adjustment (adj.); #Baseline body-mass index, systolic and diastolic blood pressure, hypertension, diabetes, active smoking, and a history of myocardial infarction, heart failure, atrial fibrillation / flutter, supraventricular tachycardias, myocardial infarction, device implantation, beta-blockers and calcium channel-blockers; $Lipid covariates including; Low Density Lipoprotein (LDL), High Density Lipoproteins (HDL) Total Cholesterol and Triglycerides; @Red blood cell covariates including; Red Blood Cell Count (RBC), Packed Cell Volume (PCV), Mean Corpuscular Volume (MCV) and Hemoglobin count (Hb).

evidence for directional (unbalanced) pleiotropy (Table 4). When we used the genetic variant coefficients derived from the associations with resting heart rate when restricted to healthy individuals (Table 1), the prediction of all-cause mortality remained similar (Table 4), further supporting the idea that underlying diseases or heart-rate-lowering medication did not confound our observation. The association with all-cause mortality also persisted when we used genetic variant coefficients estimated in the replication sample. When extrapolating the estimates from our sensitivity analyses, (ranging from

1.11 to 1.29 (Table 4)), this would translate to a reduction in life expectancy for males between 1.9 up to 4.1 years and females 1.8 up to 3.7 years per 5 bpm increase in resting heart rate.

At 19 of our 64 loci the sentinel genetic variant or a genetic variant in linkage disequilibrium (LD; $r^2 > 0.8$) have reported GWAS associations. These include lipid, metabolic and blood pressure related traits (Supplementary Table 7). The 64 loci were highly enriched for deoxyribonuclease I (DNase I) hypersensitive sites, marking transcriptionally active regions of the genome in human fetal heart tissue (Figure 2a). Enrichment testing of expression in 209 tissue and cell types identified cardiovascular tissues and the adrenal gland to be the most relevant for our association findings (Figure 2b, Supplementary Table 8). Across the 64 loci, 1,668 annotated genes are located within 1 Mb of all the sentinel genetic variants. On the basis of proximity, the presence of non-synonymous genetic variants in high LD, *cis*-expression quantitative trait loci (eQTL) and Data-driven Expression-Prioritized Integration for Complex Traits (DEPICT)[24] analyses we prioritized 102 potential candidate genes at our 64 loci (Supplementary Note and Supplementary Tables 9-11). A systematic search

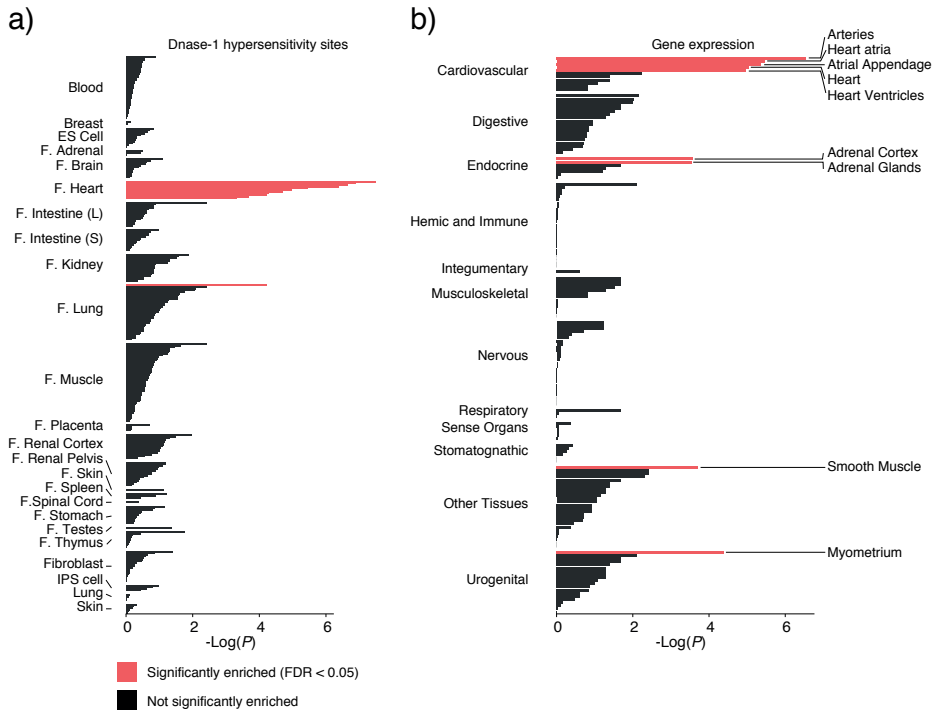


Figure 2. Biological insights
(a) The 64 genomewide associated variants were enriched within DHSs of fetal heart tissue (N=8) specifically, suggesting that functionality of regulatory DNA elements may underlie some of the associations. (b) DEPICT identified statistically significant enrichment for 9 tissue annotations of which cardiovascular tissues were the most relevant for the heart rate associated loci.

of our 102 candidate genes in Online Mendelian Inheritance in Man (OMIM) identified several Mendelian diseases with cardiac phenotypes. These were related to cardiomyopathies (*TTN, DES, SCN5A, PLN, MYH6, MYH7* and *SPEG*), Brugada syndrome (*SCN5A, CACNA1C* and *HCN4*), long QT (*SCN5A, KCNJ5* and *ALG10*), arrhythmias (*SCN5A, HCN4, CACNA1D* and *MYH6*) and congenital heart disease (*NKX2-5, PLN, TBX20, MYH6* and *MYH7*). The DEPICT tool identified 622 significantly (false discover rate (FDR)<5%) enriched gene sets (Supplementary Tables 12 and 13). We clustered them on the basis of the correlation between scores for all genes (Supplementary Note), which resulted in 74 gene sets relevant to cardiac biology (Supplementary Figure 7).

DISCUSSION

This work highlights the unprecedented opportunities provided by large scale projects such as UK Biobank, the 100,000 genomes[25], and the Precision Medicine Initiative[26] to discover novel genetic associations and to study links with outcomes and mortality. In this GWAS and replication study, performed in 265,046 individuals, we found 46 novel genetic loci associated with resting heart rate increasing the total number of heart rate loci to 67[23]. Several epidemiologic studies have reported an association between higher resting heart rate and increased mortality from both cardiovascular and non-cardiovascular causes[3-8]. In all of these studies, this association is potentially confounded by differences in demographics and physiological characteristics such as body-mass index, smoking, alcohol consumption and blood pressure. Further, data from intervention trials do not provide a consistent link between heart rate reduction and improvement of clinical outcomes. Selective sinus-node inhibition with ivabradine has beneficial effects on outcomes in patients with chronic heart failure[21] but did not improve outcomes in patients with CAD[20].

In the present work we show that genetic variants associated with higher resting heart rate confer a risk for all-cause mortality. We studied the strength of these genetic variants with mortality and studied the role of heart rate in comparison of other, potentially confounding variables closely associated with heart rate. The genetic variants identified to be associated with heart rate were also associated with potential measured (body mass index, systolic and diastolic blood pressure, hypertension, smoking, supraventricular tachycardia and device implantation) and unmeasured confounders. However, also our analyses adjusting for covariates, allowing genetic variants to have pleiotropic effects, removing genetic variants associated with other traits, or using estimates derived from healthy participants and 130,795 independent participants consistently suggest that heart rate is linked to mortality and, by extension, to life expectancy. Indeed, only heart rate itself attenuated the association of the genetic variants with the outcome to the null. This leaves two likely possibilities: either the genetic variants exert their effect

on mortality directly via heart rate as a mediator, or the genetic variants share underlying biology, resulting in increases in heart rate as well as mortality risk. Although direct specific intervention (sinus-node inhibition) on heart rate does not consistently result in reduction in mortality[20,21,27] we hypothesize the association originates from a shared biology not targeted by sinus-node inhibition. This could involve basic cellular biology behind heart rate and, possibly, vulnerability to cardiac arrhythmias causing (sudden) death, which might contribute to all classifications of death and might eventually also be relevant for a plethora of non-cardiac diseases and conditions. This theory can be supported by the identification of predominant cardiac candidate genes at the identified loci and the colocalization of DNase hypersensitivity sites in cardiac tissue. However, alternative speculations involving basic metabolic rate, energetics, free radicals, could result in cumulative general damage and affect life span[28].

In addition to an interpretation of causation, there are several other limitations of our study that are important to acknowledge. Although recent studies[29,30] and empirical estimates on the UK10K[31] and 1000 Genomes project[32] support the use of a genome-wide significant threshold at the level of $P<5.0\times10^{-8}$, the adequacy of this value for UK Biobank has not been fully investigated. In addition, among the loci identified, a number of candidate genes have a known function relevant for cardiac conditions; however, for none of the genes have we proven it is the mechanism for the association with heart rate or all-cause mortality. Our findings are based on statistical analyses of large datasets and do not include experimental validation of each locus to identify the underlying biological mechanisms. As with all bioinformatics analyses, the results should be interpreted as hypothesis generating and requiring wet lab validation. In addition, the list of candidate genes provides only a first interpretation with arbitrarily defined guidelines used in the GWAS community to suggest genes for further evaluation. Heart rate is a complex trait, and the principal reason for genes to be associated does not necessarily imply a role via the cardiac pacemaker or sinus-node function. Owing to the relative short follow-up currently available and limited number of events, our analyses focused on all-cause mortality and a crude subdivision according to the tenth revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10). On the basis of gene and pathway analyses, differences in death due to the ICD-10 category 'circulatory system' might be expected to account for the association with all-cause mortality, but this was not observed. The reason that no association was observed with 'circulatory system' remains unknown, but it might be due to heterogenic causes of death within each category; deaths in other categories might be influenced by the heart but not attributed to it. As more subjects are genotyped and long-term follow-up data become available, future analyses may allow further differentiation within each ICD-10 category to study associations of resting heart rate with specific causes of deaths.

In conclusion, in this GWAS we identified 46 novel loci associated with resting heart rate. The loci identified as influencing resting heart rate are also implicated in overall mortality (and, consequently, life expectancy) and therefore warrant further research into the underlying mechanisms.

Data Availability Statement

The GWAS discovery data that support the findings of this study are available at, http://www.cardiomics.net.

REFERENCES

1. Schmidt-Nielsen, K. Animal Physiology: Adaptation and Environment. *New York: Cambridge University press*, 133 (1975).
2. Levine, H.J. Rest heart rate and life expectancy. *J Am Coll Cardiol* 30, 1104-6 (1997).
3. Dyer, A.R. et al. Heart rate as a prognostic factor for coronary heart disease and mortality: findings in three Chicago epidemiologic studies. *Am J Epidemiol* 112, 736-49 (1980).
4. Kannel, W.B., Kannel, C., Paffenbarger, R.S., Jr. & Cupples, L.A. Heart rate and cardiovascular mortality: the Framingham Study. *Am Heart J* 113, 1489-94 (1987).
5. Gillum, R.F., Makuc, D.M. & Feldman, J.J. Pulse rate, coronary heart disease, and death: the NHANES I Epidemiologic Follow-up Study. *Am Heart J* 121, 172-7 (1991).
6. Greenland, P. et al. Resting heart rate is a risk factor for cardiovascular and noncardiovascular mortality: the Chicago Heart Association Detection Project in Industry. *Am J Epidemiol* 149, 853-62 (1999).
7. Kristal-Boneh, E., Silber, H., Harari, G. & Froom, P. The association of resting heart rate with cardiovascular, cancer and all-cause mortality. Eight year follow-up of 3527 male Israeli employees (the CORDIS Study). *Eur Heart J* 21, 116-24 (2000).
8. Reunanen, A. et al. Heart rate and mortality. *J Intern Med* 247, 231-9 (2000).
9. Kolloch, R. et al. Impact of resting heart rate on outcomes in hypertensive patients with coronary artery disease: findings from the INternational VErapamil-SR/trandolapril STudy (INVEST). *Eur Heart J* 29, 1327-34 (2008).
10. Diaz, A., Bourassa, M.G., Guertin, M.C. & Tardif, J.C. Long-term prognostic value of resting heart rate in patients with suspected or proven coronary artery disease. *Eur Heart J* 26, 967-74 (2005).
11. Bohm, M. et al. Heart rate as a risk factor in chronic heart failure (SHIFT): the association between heart rate and outcomes in a randomised placebo-controlled trial. *Lancet* 376, 886-94 (2010).
12. Grassi, G. et al. Heart rate as marker of sympathetic activity. *J Hypertens* 16, 1635-9 (1998).
13. Bohm, M., Reil, J.C., Deedwania, P., Kim, J.B. & Borer, J.S. Resting heart rate: risk indicator and emerging risk factor in cardiovascular disease. *Am J Med* 128, 219-28 (2015).
14. Aladin, A.I. et al. The Association of Resting Heart Rate and Incident Hypertension: The Henry Ford Hospital Exercise Testing (FIT) Project. *Am J Hypertens* (2015).
15. Jiang, X. et al. Metabolic syndrome is associated with and predicted by resting heart rate: a cross-sectional and longitudinal study. *Heart* 101, 44-9 (2015).
16. Caetano, J. & Delgado Alves, J. Heart rate and cardiovascular protection. *Eur J Intern Med* 26, 217-22 (2015).
17. Bangalore, S. et al. beta-Blocker use and clinical outcomes in stable outpatients with and without coronary artery disease. *JAMA* 308, 1340-9 (2012).
18. Messerli, F.H., Grossman, E. & Goldbourt, U. Are beta-blockers efficacious as first-line therapy for hypertension in the elderly? A systematic review. *JAMA* 279, 1903-7 (1998).
19. Van Gelder, I.C. et al. Lenient versus strict rate control in patients with atrial fibrillation. *N Engl J Med* 362, 1363-73 (2010).
20. Fox, K. et al. Ivabradine in stable coronary artery disease without clinical heart failure. *N Engl J Med* 371, 1091-9 (2014).
21. Swedberg, K. et al. Ivabradine and outcomes in chronic heart failure (SHIFT): a randomised placebo-controlled study. *Lancet* 376, 875-85 (2010).
22. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 12, e1001779 (2015).

23. den Hoed, M. et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat Genet* 45, 621-31 (2013).
24. Pers, T.H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* 6, 5890 (2015).
25. Siva, N. UK gears up to decode 100,000 genomes from NHS patients. *Lancet* 385, 103-4 (2015).
26. Collins, F.S. & Varmus, H. A new initiative on precision medicine. *N Engl J Med* 372, 793-5 (2015).
27. Fox, K., Ford, I., Steg, P.G., Tendera, M. & Ferrari, R. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): a randomised, double-blind, placebo-controlled trial. *Lancet* 372, 807-16 (2008).
28. Azbel, M. Universal biological scaling and mortality. *Proc Natl Acad Sci U S A* 91, 12453-7 (1994).
29. Davies, G. et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol Psychiatry* 21, 758-67 (2016).
30. Lane, J.M. et al. Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. *Nat Commun* 7, 10889 (2016).
31. Xu, C. et al. Estimating genome-wide significance for whole-genome sequencing studies. *Genet Epidemiol* 38, 281-90 (2014).
32. Kanai, M., Tanaka, T. & Okada, Y. Empirical estimation of genome-wide significance thresholds based on the 1000 Genomes Project data set. *J Hum Genet* (2016).

METHODS

Populations

Discovery: To identify genetic variants associated with resting heart rate we analyzed 134,251 participants from the UK Biobank. The UK Biobank recruited persons aged 40 - 69 years who were registered with a general medical practitioner within the UK National Health Service (NHS). In total, the study recruited 503,325 individuals between 2006 and 2010. The study has approval from the North West Multi-centre Research Ethics Committee, and all participants provided informed consent. Detailed methods used by UK Biobank have been described elsewhere[22]. For sensitivity analyses we defined a subgroup of healthy individuals which were free of any (prevalent or incident) disease(s) and diagnosis and confirmed they were not using heart rate modifying medication (beta-blockers, and calcium-channel blockers drugs (N=11,405)).

Replication: Replication of genome wide significant lead SNPs was undertaken in the meta-analysed data of 130,795 individuals derived from 23andMe, deCODE, PREVEND and LifeLines sample collections (Supplementary Table 14).

Ascertainment of resting heart rate

As detailed in the Supplementary Note, resting heart rate in UK Biobank was assessed by two methods: an automated reading during blood pressure measurement (in 501,340 participants) and during arterial stiffness measurement using the pulse waveform obtained of the finger with an infrared sensor (in 170,790 participants). Multiple available measurements for one individual were averaged.

Ascertainment of cardiovascular events and mortality

The prevalence and incidence of cardiovascular risk factors (Supplementary Table 15), conditions and events in UK Biobank were captured through data collected at the Assessment Centre in-patient Health Episode Statistics (HES) as detailed in the Supplementary Note. Information on the cause of death was obtained via the National Health Service (NHS) Information Centre for participants from England and Wales, and from the NHS Central Register, Scotland for participants from Scotland. All-cause mortality included all deaths occurring before February 17th 2014 (or December 31st 2012, for the participants enrolled in Scotland).

Genotyping and imputation

Genotype imputation data in UK Biobank was available for 152,249 (25%) individuals as of May 2015 [Interim Data Release]. In 49,923 individuals genotyping was performed as part of the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE; 807,411 variants) project and in an additional 102,326 individuals genotyping was performed on the UK

Biobank Axiom array (Affymetrix; 820,967 variants). Imputed genotype data was provided by UK Biobank based on merged UK10K and 1000 Genomes Phase 3 panel produced by the Wellcome Trust Centre for Human, resulting in 72,355,667 single nucleotide polymorphisms, short indels and large structural variants. Quality control for genotyping has been performed prior to analysis and described in detail elsewhere[33]. We excluded variants with minor allele frequency of <0.001, and information measure <0.3 leaving 19,941,912 variants for the current analyses. Samples were excluded from our analyses if they had at least one related sample (N=17,308) on the basis of genetic-relatedness factor data and high missingness or excess heterozygosity (N=480). A flow diagram of samples sizes after exclusion of participants is provided in Supplementary Figure 8.

Statistical analysis

A genome-wide association study (GWAS) was performed using SNPTEST with 19,941,912 genotyped or imputed genetic variants and resting heart rate in 134,251 individuals of UK Biobank using linear regression assuming an additive genetic model. Covariates included in the model were: age, age^2, sex, the first 10 principal components, and genotyping array. Independent genetic loci were defined as 1Mb at either side of the genetic variant that showed the strongest association in a given locus and pair-wise LD r^2<0.1. The strongest associated variant (lowest *P*-value) within a locus with at least one genetic variant at $P<5\times10^{-8}$ was designated the sentinel genetic variant. Replication of these variants was undertaken in the 23andMe, deCODE, Prevend and LifeLines cohorts using fixed-effects meta-analysis by inverse variance weighting (Supplementary Table 14). An association was considered replicated if (1) the direction of effect was concordant, (2) the replication-*P*<0.025 (one-way), and (3) meta-$P<5\times10^{-8}$. For detecting secondary associations not explained by the sentinel genetic variant at each locus, we repeated the GWAS while including all sentinel genetic variants ($P<5\times10^{-8}$) as covariates in a conditional analysis. Potential modifier effects of gender, β-adrenergic receptor-blocking agent (beta-blockers), and calcium-channels blockers drugs on resting heart rate were assessed by an interaction test (Bonferroni adjusted for the number (*n*) of tests ($P<0.05/n$)).

We used genetic variants as instrumental variables to study the relationship of resting heart rate with outcomes (Mendelian Randomization). To this end we defined a larger set of independent loci at the previously specified hypothesis-generating threshold ($P<1\times10^{-5}$) to increase power[34,35]. For our main analysis we calculated the putative association between resting heart rate (per 5 bpm) and outcome mediated through that variant (β_3 values) from the direct measurements of the effect size of the association between the variant and resting heart rate (β_1) and the effect size of the association between the variant and outcome (β_2), as described previously[37]. The value of β_3 can be interpreted as the hazard ratio for outcome per 5 bpm increase in genetically de-

termined resting heart rate. Inverse-variance-weighted random-effects meta-analysis was used to combine individual β_3 estimates providing additional power to assess the overall association between genetically determined resting heart rate and mortality. Cochran's Q statistic was used to assess heterogeneity among β_3 estimates. We also created a weighted genetic risk score (GRS) by first multiplying for each individual the effect size of the association between the variant and resting heart rate (β_1) with the number of alleles 0-2 of each genetic variant and then summing all products. An unweighted GRS was created by summing the number of resting heart rate-increasing alleles 0-2 of each associated genetic variant.

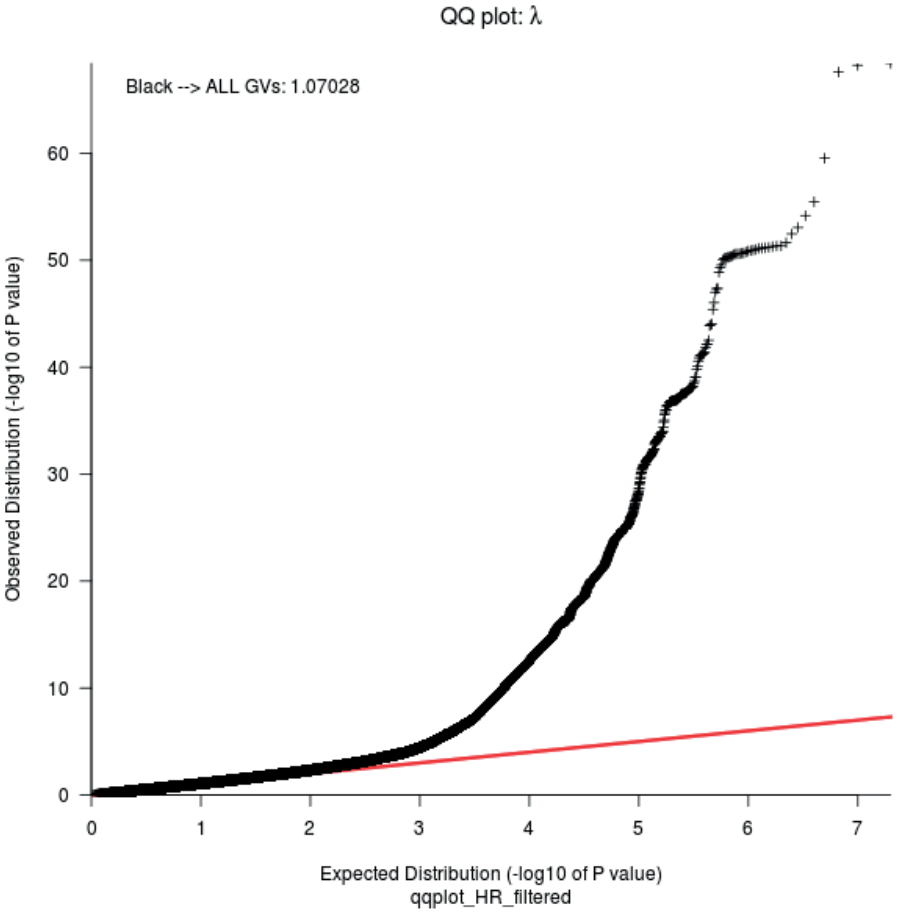
To examine the robustness of our findings as well as the possibility of pleiotropic or other confounding and mediation effects we included covariates and the phenotype resting heart rate into the Cox regression models. We excluded all genetic variants that were also individually nominally associated ($P<0.05$) with covariates, and performed multivariable Mendelian randomization[37] to account for variables not available in UK Biobank, and used the MR-Egger regression method to test for evidence of pleiotropy[38] (details provided in Supplementary Note and Supplementary Figures 9 and 10). As an alternative strategy to exclude confounding due to prevalent disease or medication use, we estimated the associations of each genetic variant with resting heart rate (β_1) in the subgroup of 11,405 healthy individuals (defined above) to calculate the hazard ratio for outcome. We estimated the impact on life expectancy using the National Life Tables of the United Kingdom provided by the Office of National Statistics (ONS; www.ons.gov.uk) of 2011-2013 separately for males and females (Supplementary Note). Details of analyses performed to gain insights in the biological pathways and tissues underlying the genome-wide significant loci are provided in the Supplementary Note.

METHODS-ONLY REFERENCES

33. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource Information for researchers. *Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-web.pdf*.
34. Purcell, S.M. et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460, 748-52 (2009).
35. Thanassoulis, G. et al. Genetic associations with valvular calcification and aortic stenosis. *N Engl J Med* 368, 503-12 (2013).
36. Nelson, C.P. et al. Genetically determined height and coronary artery disease. *N Engl J Med* 372, 1608-18 (2015).
37. Burgess, S. & Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 181, 251-60 (2015).
38. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 44, 512-25 (2015).

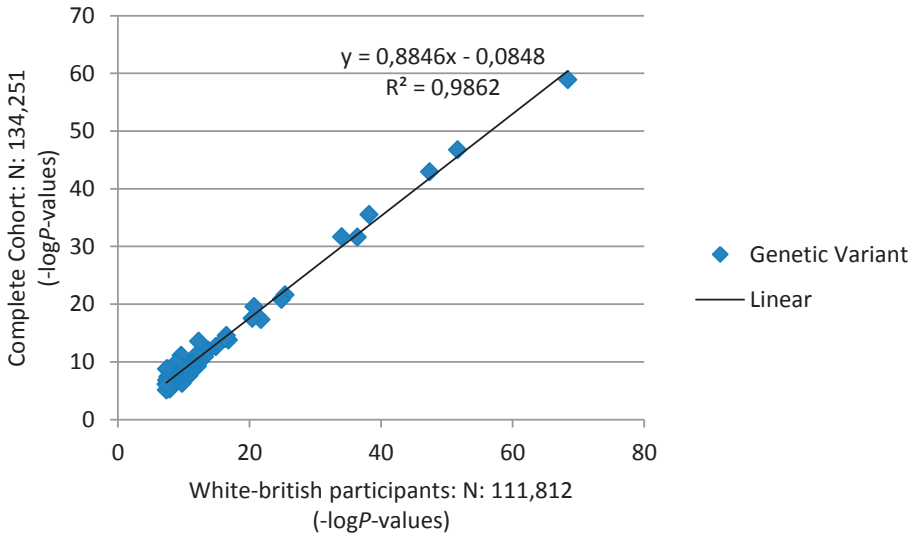
SUPPLEMENTARY APPENDIX

Supplementary Figure 1 | Quantile-quantile plot of genetic variants after discovery-analysis 43
Supplementary Figure 2 | Regional association plots for of GWAS analysis on resting heart rate 44
Supplementary Figure 3 | Scatterplot of lead genetic variants *P*-values from the Complete cohort versus White-British participants 45
Supplementary Figure 4 | Forest Plot showing the effect size of resting heart rate on the risk of all-cause mortality for heart rate-associated genetic variant 46
Supplementary Figure 5 | Kaplan-Meier failure curve death weighted GRS 47
Supplementary Figure 6 | Interpreting the association between genetically determined heart rate and all-cause mortality 48
Supplementary Figure 7 | Biological insights 49
Supplementary Figure 8 | Sample selection strategy 50
Supplementary Figure 9 | Scatter plot of genetic associations with all-cause mortality 51
Supplementary Figure 10 | Funnel plot of instrument variable strength 52
Supplementary Note 53
Ascertainment of resting heart rate 53
Estimate of heritability 53
Ascertainment of mortality 53
Definitions used UK Biobank 53
Biological pathway and enrichment analyses 54
Prioritization of potential candidate genes 55
Associations between genetically determined heart rate and all-cause mortality 55
Primary analysis: 55
Sensitivity analyses: 56
Genetic Risk Score (GRS): 56
Correct for risk-factors in Cox regression model: 56
Excluding variants that showed an association with risk-factors: 56
Multivariable regression of beta-coefficients method: 56
MR-Egger test: 57
Re-estimate betas using healthy individuals: 57
Re-estimate betas using replication cohort: 57
Calculating genetic correlation between resting heart rate and cardiovascular disease traits using BOLT-REML: 57
Sort individuals with a mortality event by primary cause of death above or below the median value calculated GRS: 57
Estimation of life expectancy 58
Supplementary References 59



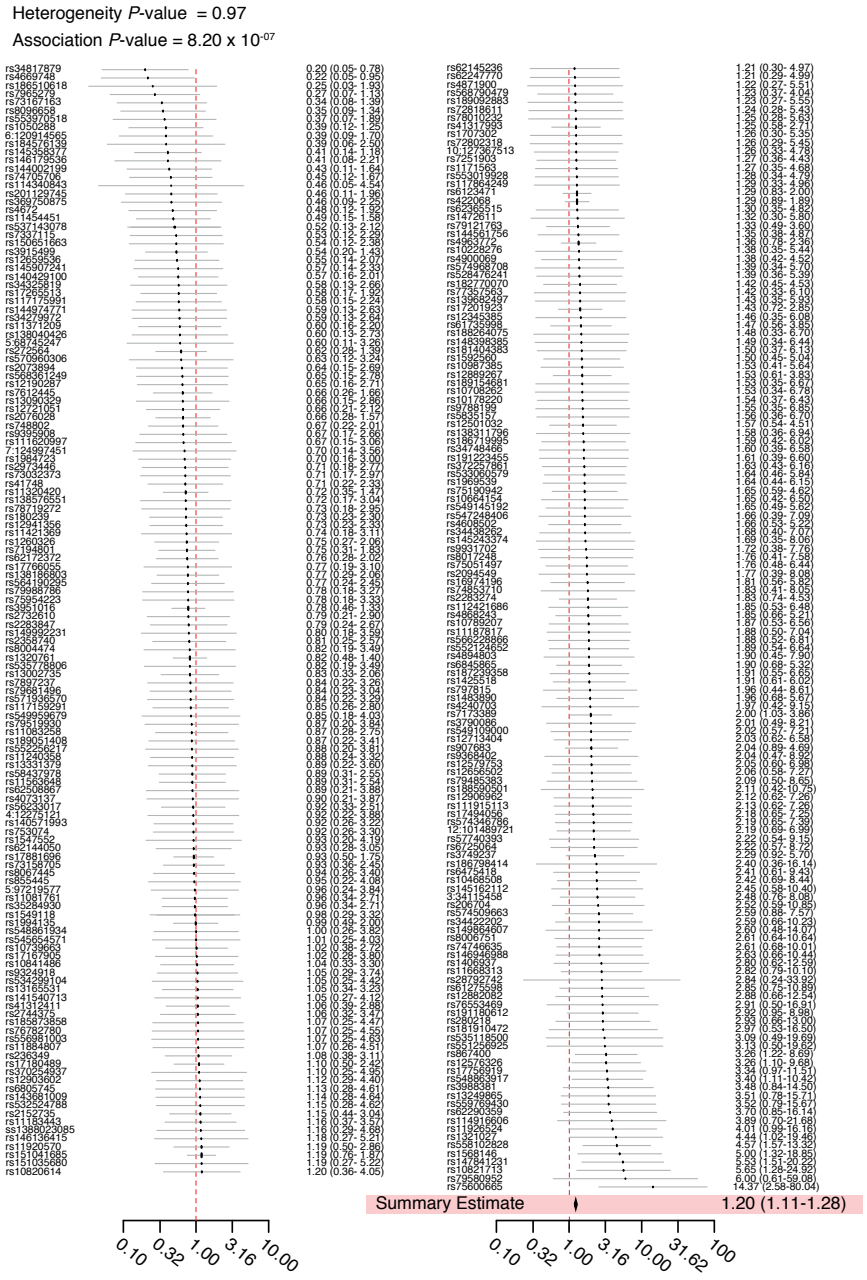
Supplementary Figure 1. Quantile-quantile plot of genetic variants after discovery-analysis
Data is shown for the discovery analysis of N:134,251 individuals excluding genetic variants with minor allele frequency of <0.001, and information measure <0.3 leaving 19,941,912 genetic variants.

Supplementary Figure 2. Regional association plots for of GWAS analysis on resting heart rate
Regional plots for the resting heart rate phenotype sentinel genetic variants. At each region pairwise LD with the sentinel genetic variant is indicated.
See separate pdf-document (Supplementary Figure 2.pdf) online



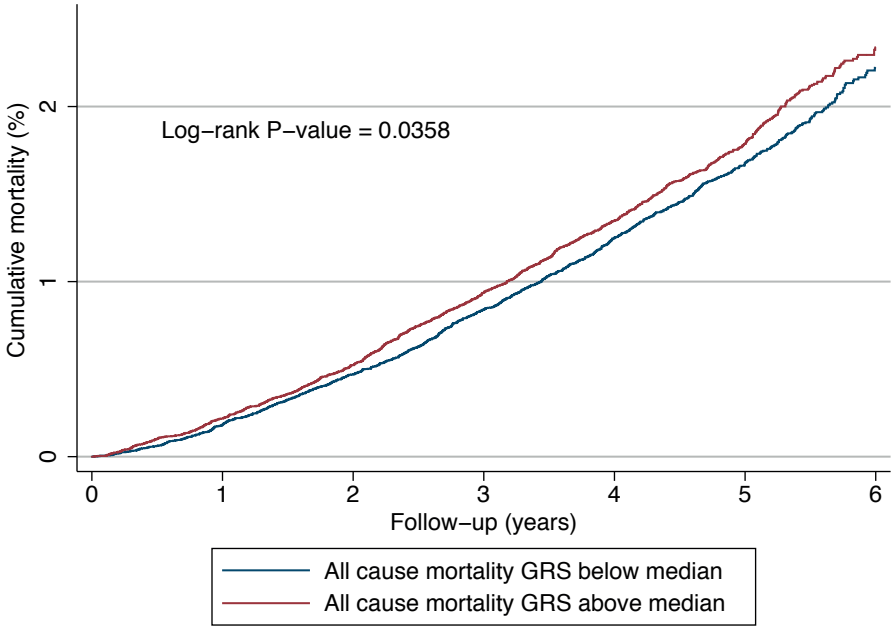
Supplementary Figure 3. Scatterplot of lead genetic variants P-values from the Complete cohort versus White-British participants

Scatter plot of lead genetic variant association P-values with resting heart rate in the complete cohort N:134,251 (y-axis) against lead genetic variants association P-values with resting heart rate in the White-British participants N:111,812 (x-axis).

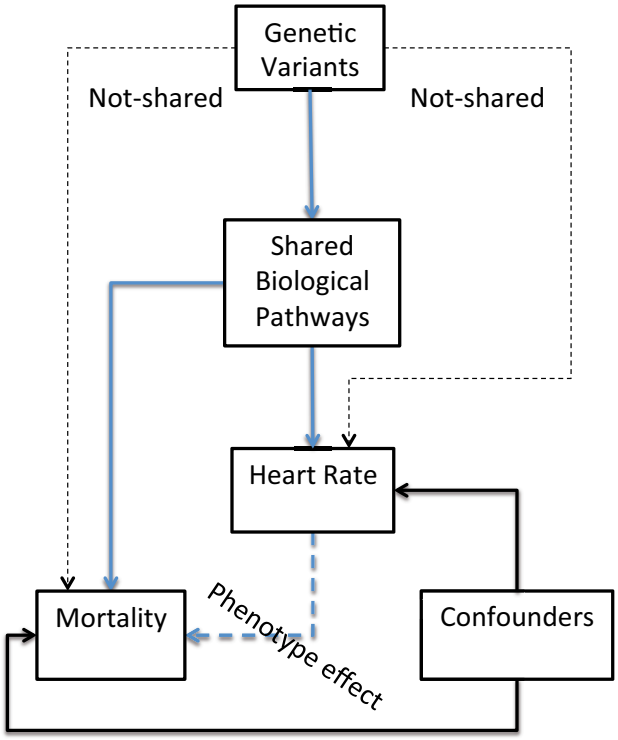


Supplementary Figure 4. Forest Plot showing the effect size of resting heart rate on the risk of all-cause mortality for heart rate-associated genetic variant

The odds ratios for each resting heart rate associated genetic variant for β3 values (i.e., the putative association between resting heart rate and all-cause mortality mediated through that variant) are shown. The β_3 odds ratios are organized in ascending values. The overall β3 estimate (shown in red) is from a random-effects meta-analysis of all genetic variants.

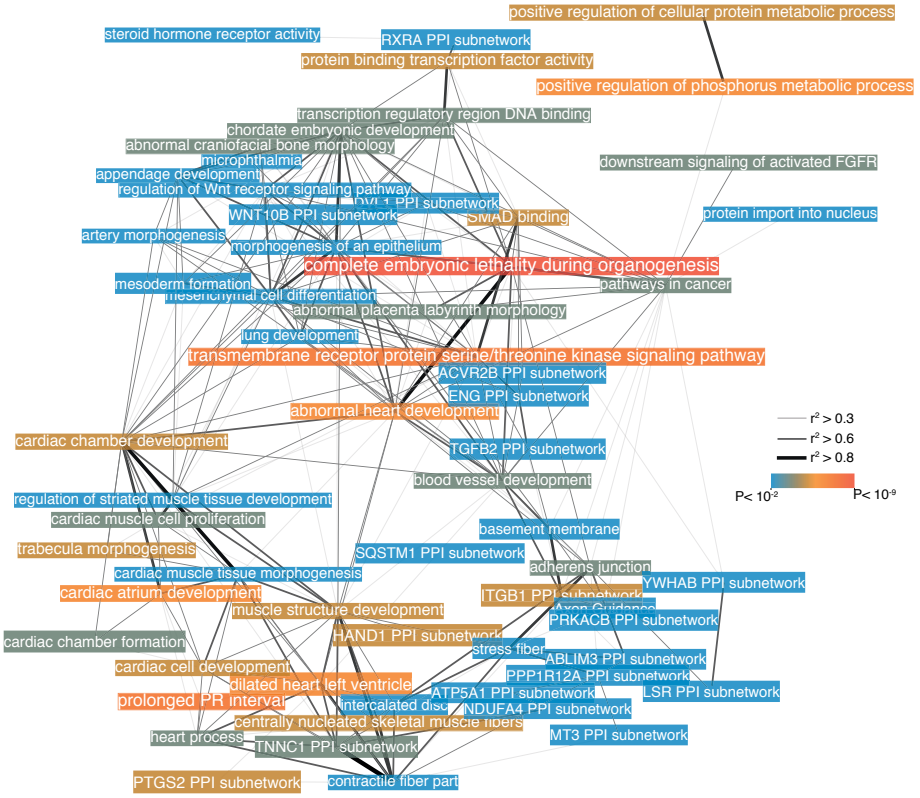


Supplementary Figure 5. Kaplan-Meier failure curve death weighted GRS
Shown are cumulative all-cause mortality in (%) divided by individuals below and above the median of the weighted Genetic Risk Score (GRS).



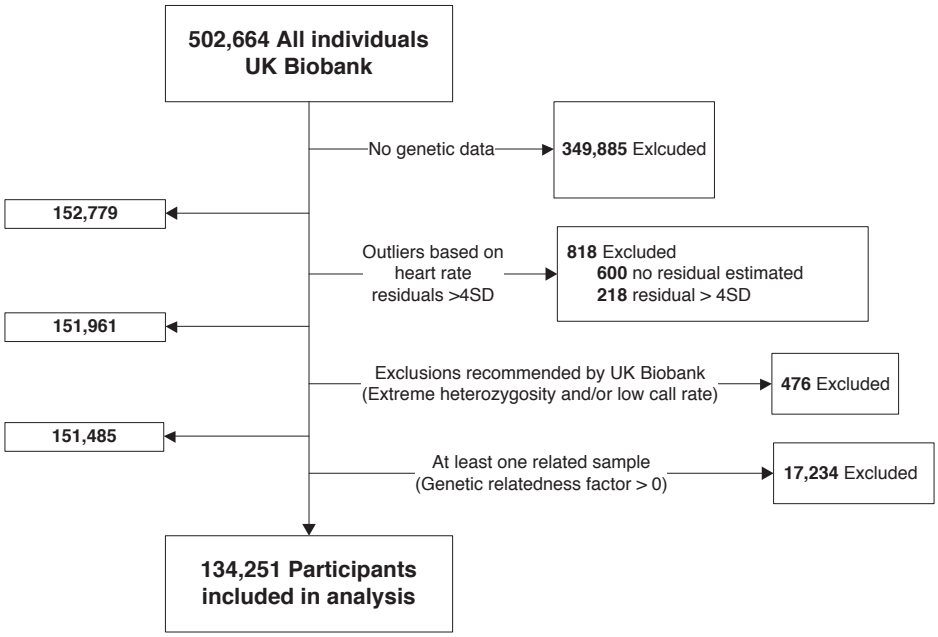
Supplementary Figure 6. Interpreting the association between genetically determined heart rate and all-cause mortality

Conceptual diagram to interpret the Association between Genetically Determined Heart Rate and all-cause Mortality. The benefits a genetic approach have been explained before[1] and essentially reduces the probability of known and unknown demographic, socioeconomic, behavioral or lifestyle confounders. These confounders have an independent effect on resting heart rate and the risk on all-cause mortality (solid black lines) and could give rise to a false association between the two factors. There is a possibility that the association between the studied genetic variants and resting heart rate and the association with mortality are through completely different mechanisms (dashed black lines). What seems to be a more possible situation on the basis of our outcomes is that resting heart rate variants affect biologic pathways, which on the one hand influence resting heart rate and on the other hand influence all-cause mortality (solid blue lines). Another possibility is that genetically predicted resting heart rate itself alters behavior or lifestyle, which then affects the risk of all-cause mortality (dashed blue line).

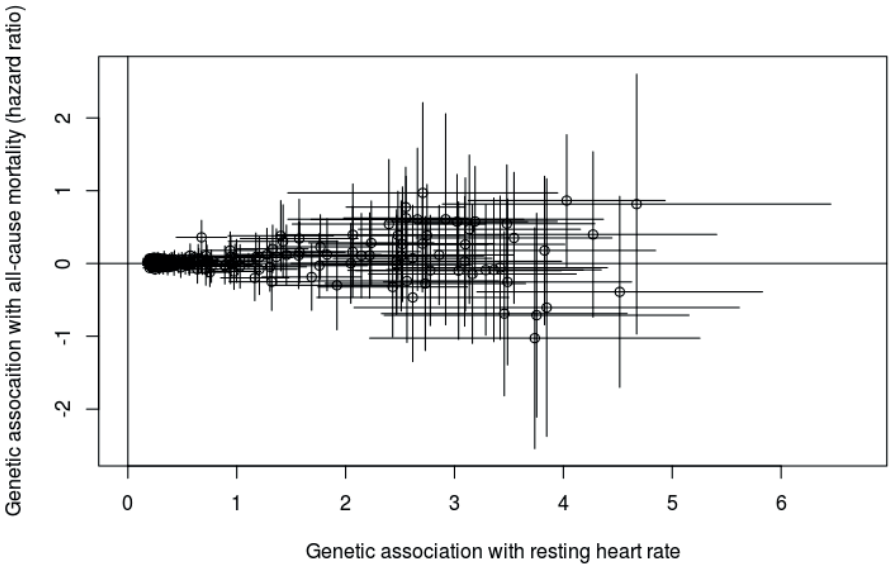


Supplementary Figure 7. Biological insights

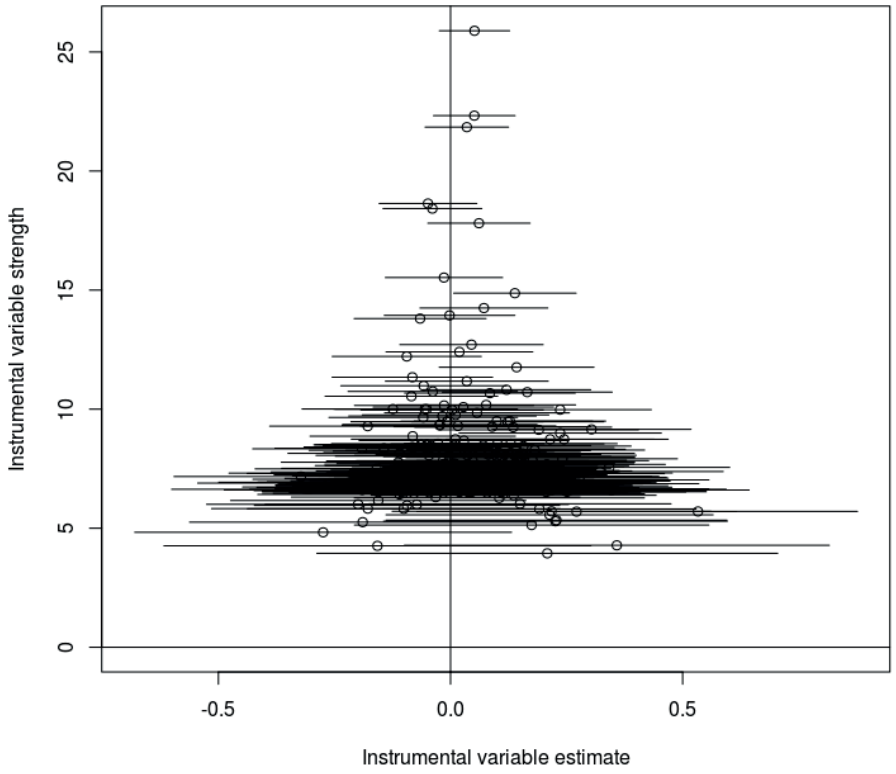
DEPICT pathway analysis identified 623 significantly enriched gene-sets relevant for heart rate. The 74 meta-gene set clusters are shown.



Supplementary Figure 8. Sample selection strategy
Scatter plot of genetic associations with all-cause mortality (y-axis) against associations with resting heart rate (x-axis) for variants associated with resting heart rate at $P<10^{-5}$ (lines represent 95% confidence intervals).



Supplementary Figure 9. Scatter plot of genetic associations with all-cause mortality
Scatter plot of genetic associations with all-cause mortality (y-axis) against associations with resting heart rate (x-axis) for variants associated with resting heart rate at $P<10^{-5}$ (lines represent 95% confidence intervals).



Supplementary Figure 10. Funnel plot of instrument variable strength

Funnel plot of instrument variable strength (y-axis, larger values indicate stronger instruments) against instrumental variable estimates for each genetic variant separately in Mendelian randomization analysis of resting heart rate on all-cause mortality using genetic variant using genetic variants throughout the genome that have been associated with resting heart rate at $P<10^{-5}$. Horizontal lines represent 95% confidence interval for the instrumental variable estimates. Solid vertical line is at the nul.

SUPPLEMENTARY NOTE

Ascertainment of resting heart rate

Resting heart rate was assessed by two methods: pulse rate - using an automated reading during blood pressure measurement field, Field-ID 102 and Field-ID 95 (in 501,340 participants); pulse rate - during arterial stiffness measurement using the pulse waveform obtained of the finger with an infra-red sensor Field-ID 4194 (in 170,790 participants). When multiple heart rate measurements were available during the first visit for an individual these measurements were averaged. In 99.7% of participants at least one single measurement was available. There were 490 participants with 1 measurement, 330,419 with 2 measurements and 170,594 with 3 heart rate measurements. The coefficient of variation derived from multiple measurements was 4.9%. We excluded individuals with extreme (>4SD) values (N=818).

Estimate of heritability

To estimate the heritability of heart rate explained by genetic variants, we employed BOLT-REML[2]. From the imputed data set (to ensure 100% call rate) all directly genotyped variants that passed quality control were extracted and pruned on LD (R^2<0.05) to obtain roughly 500k variants, as recommended[2]. Resting heart rate was adjusted for gender, age, age^2, principle components and the genotyping chip. We determined the heritability of heart rate explained by all the genetic variation which was robustly genotyped (n=514,523) to be h^2_{GWAS}=21.2% (SE=0.2%).

Ascertainment of mortality

Participant follow-up started at inclusion in the UK Biobank study and follow-up ended on Feb 17, 2014, or death, for all participants. This was for every participant apart from those enrolled in Scotland, which had complete information up to Dec 31, 2012. All-cause mortality included all deaths occurring before Feb 17, 2014. Or for the participants enrolled in Scotland Dec 31, 2012. From the National Health Service (NHS) Information Centre for participants the information about cause of death was obtained for participants from England and Wales, and from the NHS Central Register, Scotland for participants from Scotland. Detailed information about the linkage procedure is available online (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=115559).

Definitions used UK Biobank

Prevalent and incident hypertension, diabetes, myocardial infarction, heart failure, atrial fibrillation and supraventricular tachycardia was derived from self-reported (touchscreen questionnaire and verbal interview) and/or the diagnosis was captured using the Hospital Episode Statistics (HES) records (using the following ICD codes: I10, I11, I12, I13,

I14, I15 for hypertension; E10, E11, E12, E13, E14 for diabetes; I21, I22 for myocardial infarction; I42, I50 for heart failure ; I48 for atrial fibrillation; I471 for supraventricular tachycardia). The "Spell and Episode" category contains data relating to diagnoses made during hospital in-patient stay. It includes main and secondary diagnoses, coded according to the International Classification of Diseases (ICD). The main diagnosis is taken to be the main reason for the hospital admission, while secondary diagnoses are more often contributory or underlying conditions. We used both the main and secondary diagnoses for recording prevalent and incident risk factors, conditions and events.

Information on smoking status was collected using the touchscreen questionnaire at baseline visit. Information regarding device implantation (pacemaker or implantable cardioverter defibrillator) was captured through data collected at the Assessment Centre (verbal interview) HES records which are coded according to the Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4 (OPCS-4). Device implantation according the OPCS-4 was defined K59, K60, K61 and U31. Medication usage was collected during the baseline visit during a verbal interview by a trained nurse on prescription medications (Field ID 20003) this data was used for the interaction analysis of beta-blockers and calcium channel blockers. Body mass index was calculated using BMI value constructed from height and weight measured during the initial Assessment Centre visit (Field ID 21001) and Body composition estimation by impedance measurement (Field ID 23104). Blood pressure was measured using the manual reading (Field ID 93, 94) and automated reading (Field ID 4079, 4080) measurements. When multiple measurements during firs visit were available the mean of all measurements were averaged and used in the analysis.

Biological pathway and enrichment analyses

The NHGRI-EBI catalog[3] of published genome-wide association studies was queried to identify previous associations within 1Mb of and r^2 >0.8 with loci identified by the current GWAS (18 July 2015).

DEPICT systematically identifies the most gene underlying a given associated locus, tests gene sets for enrichment in associated genetic variants, and identifies tissues and cell types in which genes from associated loci are highly expressed (see Pers et al. for a detailed description)[4]. We ran DEPICT on 376 independently associated loci of 1000 Genomes phase 3 genetic variants (association *P* values < 10^{-5}; PLINK parameters, '--clump-p1 1e-5 --clump-kb 500 --clump-r2 0.05') resulting in 236 independent, autosomal DEPICT loci containing 614 genes (the major histocompatibility complex region are by default excluded).

To obtain meta-genesets presented in figure 4C, we applied the Affinity Propagation method to group similar reconstituted gene sets. Each cluster was named according to the name of the representative gene set that was automatically identified by the Affinity

Propagation method. Pearson correlations between each gene-set were then calculated using R3.1.1 and visualized using Gephi.

Enrichment analyses for deoxyribonuclease I (DNase I) hypersensitive sites, as markers for active gene transcription, at identified loci were performed to identify the biologically relevant tissues using using FORGE[5] by using the 64 sentinel genetic variants (1000G corresponding rsIDs) as input genetic variants. In short, the 64 sentinel genetic variants were overlapped with functional DNA elements *(125 samples for ENCODE, 299 for Roadmap)*. A matched background distribution of genetic variants (matched on GC content, minor allele frequency and distance to TSS) was also overlapped with the DNA elements to obtain a null distribution. Z-score statistics was used to express enrichment, enrichments above the 99.9th percentile of the normal distribution were considered statistically significant.

Prioritization of potential candidate genes

SNPsnap[6] was used to annotate all genes that are located within 1Mb of the sentinel genetic variant. Across the 64 loci, 1,668 annotated genes are located within 1 Mb of all sentinel SNPs. Among these genes, we prioritized potential candidates using an established complementary strategy[7,8]; we chose (i.) protein coding genes nearest to the sentinel SNP, and any other genes within 10kb; (ii.) Genes containing a non-synonymous SNP in high LD (r^2>0.8) with the sentinel SNP; (iii.) Protein-coding genes with *cis*-eQTL associated with sentinel SNP; (iv) based on DEPICT analyses (see below).

For our candidate genes basic knowledge was retrieved and summarized based on querying Entrez Gene (http://www.ncbi.nlm.nih.gov), GeneCards (http://www.genecards.org), the UniProt Knowledgebase (http://www.uniprot.org) and the Online Mendelian Inheritance in Man catalog (http://www.omim.org/).

Associations between genetically determined heart rate and all-cause mortality

We applied multiple strategies and methods to examine the relationship between genetically determined resting heart rate and mortality.

Primary analysis:

Standard Mendelian analysis: for each resting heart rate associated variant we calculated β3 values (the putative association between resting heart rate (per 5 beats per minute (bpm) and outcome mediated through that variant) from the direct measurements of β1 (the effect size of the association between the variant and resting heart rate) and β2 (the effect size of the association between the variant and outcome), as described previously[1]. The value of β3 can be interpreted as the hazard ratio for outcome per 5 bpm increase in genetically determined resting heart rate. Inverse-variance-weighted random-effects meta-analysis was used to combine individual β3 estimates providing

additional power to assess the overall association between genetically determined resting heart rate and mortality. Cochran's Q statistic was used to assess heterogeneity among genetic variant estimates (Supplementary Figure 4).

Sensitivity analyses:

Genetic Risk Score (GRS):

We also created a weighted genetic risk score (GRS) by first multiplying for each individual the effect size of the association between the variant and resting heart rate (β_1) with the number of alleles 0-2 of each genetic variant and then summing all products. An unweighted GRS was created by summing the number of resting heart rate-increasing alleles 0-2 of each associated genetic variant. We assessed the association of the GRS with all cause-mortality using the Cox-regressions adjusted for age and sex.

Correct for risk-factors in Cox regression model:

To test whether the effect of our resting heart rate genetic variants on all-cause mortality is driven by cardiovascular risk factors we performed multivariable Cox regression models adjusted for age and sex (Table 4).

Excluding variants that showed an association with risk-factors:

We performed additional analyses to establish whether variants that are associated with any of the variables of Table 3 (see manuscript) might be driving the observed association with all-cause mortality. For this, we excluded variants that showed an association ($P<0.05$) with these risk factors, leaving 55 genetic variants. We re-calculated the association between resting heart rate and outcome mediated through these variants (Table 4).

Multivariable regression of beta-coefficients method:

To entangle pleiotropy from mediation we used complementary multivariable regression of beta-coefficients[9]. This method is robust to violations of the instrumental variable assumptions due to pleiotropic effects on measured variables, and can therefore entangle pleiotropy from mediation (Table 4). In addition, this method allowed us to assess whether lipids (LDL, TGL, HDL and TG) and red blood cell phenotypes (RBC, PCV, MCV and Hb) obtained through previously published GWAS lays on the causal pathway between heart rate and all-cause mortality. For lipids we downloaded data from large-scale genome-wide association meta-analyses of lipid traits of ENGAGE (European Network for Genetic and Genomic Epidemiology)[10]. In order to perform this analysis we selected SNPs from our GWAS which were in high LD ($r^2_{>}0.8$ in UK Biobank) with the reported SNPs from the previously published GWAS. We matched 173 of the genetic variants of the 272 resting heart rate genetic variants. For red blood cell phenotypes we

used data from van der Harst et al[8] and matched 209 of the genetic variants of the 272 resting heart rate genetic variants (Table 4).

MR-Egger test:
To test for evidence of pleiotropy (the phenomenon that a genetic variant directly affects multiple phenotypic traits) we applied the MR-Egger regression test. This test has been developed to test and correct for directional (unbalanced) pleiotropy[11]. We assessed whether there was heterogeneity in the causal estimates calculated using each genetic variant separately via a scatter plot and funnel (Supplementary Figure 9 and 10).

Re-estimate betas using healthy individuals:
To determine whether our GWAS on resting heart rate was driven by genetic variants of which the strength of association (Beta) was influenced by prevalent (subclinical) diseases or drug treatment, we re-analyzed estimates derived from individuals free of any disease (using questionnaire data and Hospital Episode Statistics) and heart rate modifying medication (beta-blockers, or calcium channel-blockers). We identified 11,405 (8.5%) of subjects fulfilling these stringent criteria (Table 1). We adjusted all the Beta's based on this healthy cohort and re-calculated the association between resting heart rate and outcome mediated through these variants.

Re-estimate betas using replication cohort:
To avoid any weak instrument bias, winner's curse bias and population stratification we estimated the hazard ratio on mortality by using betas estimated on 130,795 individuals from our replication meta-analysis.

Calculating genetic correlation between resting heart rate and cardiovascular disease traits using BOLT-REML:
Bivariate REML analyses was performed using BOLT-REML[2] to estimate the genetic correlation between heart rate and other traits (Supplementary Table 6). From the imputed data set (to ensure 100% call rate) all directly genotyped variants that passed quality control were extracted and pruned on LD (R^2<0.05) to obtain roughly 500k variants, as recommended[2]. Gender, age, principle components and genotyping chip we included as covariates in the analyses. Liability scale was estimated for dichotomous traits using linear transformation[12].

Sort individuals with a mortality event by primary cause of death above or below the median value calculated GRS:
We sorted individuals with a mortality event by primary cause of death subdivided into ICD-10 chapters and whether their genetically-predicted resting heart rate was above or

below the median value calculated using GRS including variants with $P<10^{-5}$ as shown in Supplementary Table 7.

Estimation of life expectancy

We calculated life expectancy estimates based on the National Life Tables of the United Kingdom provided by the Office of National Statistics (ONS; www.ons.gov.uk) of 2011-13 separately for males and females. We multiplied all the qx (is the mortality rate between age x and (x +1), that is the probability that a person aged x exact will die before reaching age (x +1)) for each age with the hazard ratio derived from our mortality analysis. We recalculated dx (is the number dying between exact age x and (x +1) described similarly to lx, that is dx=lx-lx+1.) and then approximated life-expectancy ∑Mt=xtdt/∑Mt=xdt+0.5 for each age from 0 to 90 and subtracted it from the same calculation using a hazard ratio of 1.0.

Supplemental tables (excel file) are available online.

SUPPLEMENTARY REFERENCES

1. Nelson, C.P. et al. Genetically determined height and coronary artery disease. *N Engl J Med* 372, 1608-18 (2015).
2. Loh, P.R. et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat Genet* 47, 1385-92 (2015).
3. Hindorff, L.A. et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106, 9362-7 (2009).
4. Pers, T.H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* 6, 5890 (2015).
5. Dunham, I., Kulesha, E., Morganella, S. & Birney, E. FORGE : A tool to discover cell specific enrichments of GWAS associated SNPs in regulatory regions. *submitted* (2015).
6. Pers, T.H., Timshel, P. & Hirschhorn, J.N. SNPsnap: a Web-based tool for identification and annotation of matched SNPs. *Bioinformatics* 31, 418-20 (2015).
7. Gieger, C. et al. New gene functions in megakaryopoiesis and platelet formation. *Nature* 480, 201-8 (2011).
8. van der Harst, P. et al. Seventy-five genetic loci influencing the human red blood cell. *Nature* 492, 369-75 (2012).
9. Burgess, S. & Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 181, 251-60 (2015).
10. Surakka, I. et al. The impact of low-frequency and rare variants on lipid levels. *Nat Genet* 47, 589-97 (2015).
11. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 44, 512-25 (2015).
12. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88, 294-305 (2011).

Chapter 3

Telomere length and Risk of Cardiovascular Disease and Cancer

Ruben N. Eppinga*, M. Abdullah Said*, Yanick Hagemeijer, Niek Verweij, Pim van der Harst

* both authors contributed equally

Adapted from J Am Coll Cardiol. 2017 Jul 25;70(4):506-507

Telomeres are DNA repeat structures with protein complexes capping the ends of chromosomes important for chromosomal stability and cellular integrity (1). Telomeres shorten with each cell division and under environmental conditions such as oxidative stress. Therefore, telomere length (TL) has been proposed to reflect biological age (1). Many associations between shorter TL and various age-associated cardiovascular conditions have been reported, including hypertension, coronary heart disease, and heart failure (1). Shorter and longer TL have also been linked to specific malignancies (2). These cross-sectional data do not provide evidence for causality. Here, we report a Mendelian Randomization using seven sequence variants (SVs) previously associated with TL (3) and created a genetic risk score (GRS). We studied the association of genetically determined TL (gTL) with cardiovascular conditions, and mortality in 134,773 individuals of the UK Biobank population (4).

Genotypes based on UK Biobank arrays (BiLEVE and Axiom) imputed to the merged UK10K and 1000 Genomes Phase 3 panel were used. Participant follow-up started at inclusion and ended at death or on Feb 17, 2014 (participants enrolled in Scotland Dec 31, 2012). We examined the association of both the individual SVs and GRS based on a summation of all seven SVs (weighted to their effect sizes, as reported in (3)) with hypertension, diabetes, cardiovascular disease (CVD), cancer, and mortality, as previously described (5). We adjusted our analyses for age, sex, genotyping array and the first 10 Principal Components (generated by flashPCA based on covariance) that were provided by UK Biobank. Two-sided *P*-values<0.05 were considered statistically significant.

During the 1.2 year follow-up period, 2,395 (1.8%) participants died, of which 756 (0.6%) due to CVD and 1,499 (1.1%) due to cancer. The prevalence of overall CVD was 34.9% (n=46,979). Although the effect size of individual SVs was small for CVD, the combined effect of all seven SVs was substantial. Shorter gTL was associated with 14% (95% Confidence interval [CI], 5-23%; *P*=0.004) lower risk of CVD per SD change in gTL (Figure 1A). The prevalence of hypertension was 31% (n=41,847). Two SVs, *rs10936599* (*TERC*) and *rs9420907* (*OBFC1*), were individually associated (*P*-values<0.05) with hypertension. The weighted linear combination of all seven SVs was associated with 16% (95% CI, 7-25%; *P*=0.002) decreased hypertension risk per SD shorter gTL. The prevalence of overall cancer was 16.7% (n=22,448). Four SVs were associated with overall cancer. The strongest association was observed for *rs9420907* (*OBFC1*; Figure 1B). The weighted linear combination of the SVs showed 37% (95% CI, 29-45%; *P*<0.001) decreased overall cancer risk and a 45% (95% CI, 12-65%; *P*=0.01) decreased cancer mortality risk per SD shorter gTL. We did not observe an association between gTL and diabetes (n = 7,969 [5.9%]; 9%; 95% CI: −12% to 26%; *P*=0.38) or all-cause mortality (19%; 95% CI, −17%-44%; *P*=0.26). Previous studies suggested associations between shorter TL with various CVD conditions (1). Although the exact origin of these associations remains to be elucidated, a first indication for causality was derived from a sub-analysis of CARDIoGRAM. In this previous

work, a 21% (95% CI, 5-35%) increased risk of coronary heart disease was observed per SD shorter gTL (3). We now provide independent data further supporting this causal association between longer gTL and both overall CVD and hypertension. Similar to CARDIoGRAM, *rs7675998* (NAF1) had a contrasting effect for both CVD and hypertension risk. Our results on cancer also suggest causality for previously reported associations between shorter TL and decreased cancer risk, although contrasting findings have also been reported and the reason for this discrepancy remains to be resolved (2). In conclusion, we applied a Mendelian Randomization approach and report evidence for a causal link between longer gTL with CVD, hypertension and cancer in 134,773 participants of the UK Biobank.

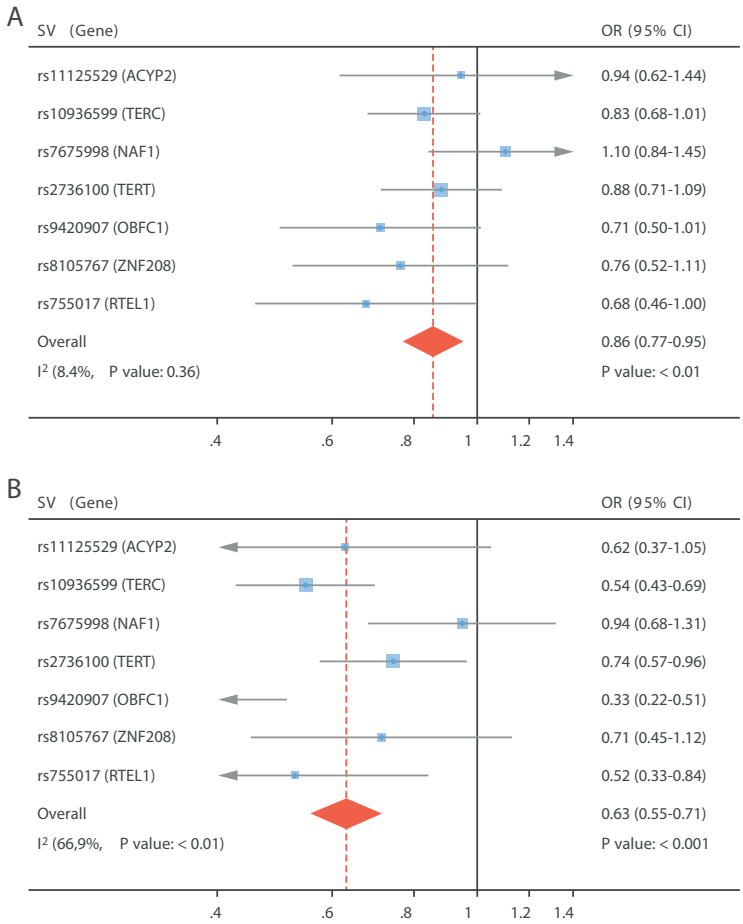


Figure 1. gTL variants and risk of CVD (A) and cancer (B)
Forest plots display the effects of shorter gTL on CVD and cancer risk for each gTL sequence variant (SV). The overall effect is from fixed-effects meta-analysis of all sequence variants. Odds ratio (OR) shown with 95% confidence interval (CI) relates to a change in risk per-SD shortening in gTL.

REFERENCES

1. Samani NJ, van der Harst P. Biological ageing and cardiovascular disease. *Heart.* 2008;94:537-9.
2. Rode L, Nordestgaard BG, Bojesen SE. Peripheral blood leukocyte telomere length and mortality among 64,637 individuals from the general population. *J Natl Cancer Inst.* 2015;107:djv074.
3. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet.* 2013;45:422,7,427e1-2.
4. Sudlow C, Gallacher J, Allen N, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12:e1001779.
5. Nelson CP, Hamby SE, Saleheen D, et al. Genetically determined height and coronary artery disease. *N Engl J Med.* 2015;372:1608-18.

Chapter 4

Identification of 15 Novel Risk Loci for Coronary Artery Disease and Genetic Risk of Recurrent Events, Atrial Fibrillation and Heart Failure

Niek Verweij, Ruben N. Eppinga, Yanick Hagemeijer, Pim van der Harst

Adapted from Sci Rep. 2017 Jun 5;7(1):2761

ABSTRACT

Coronary artery disease (CAD) is the major cause of morbidity and mortality in the world. Identification of novel genetic determinants may provide new opportunities for developing innovative strategies to predict, prevent and treat CAD. Therefore, we meta-analyzed independent genetic variants passing $p<\times10^{-5}$ in CARDIoGRAMplusC4D with novel data made available by UK Biobank. Of the 161 genetic variants studied, 71 reached genome wide significance ($p<5\times10^{-8}$) including 15 novel loci. These novel loci include multiple genes that are involved in angiogenesis (TGFB1, ITGB5, CDH13 and RHOA) and 2 independent variants in the TGFB1 locus. We also identified SGEF as a candidate gene in one of the novel CAD loci. SGEF was previously suggested as a therapeutic target based on mouse studies. The genetic risk score of CAD predicted recurrent CAD events and cardiovascular mortality. We also identified significant genetic correlations between CAD and other cardiovascular conditions, including heart failure and atrial fibrillation. In conclusion, we substantially increased the number of loci convincingly associated with CAD and provide additional biological and clinical insights.

INTRODUCTION

Coronary artery disease (CAD) is a major burden of morbidity and mortality to Western society[1]. CAD is driven by a complex interplay of multiple genetic and environmental factors that jointly give rise to a plethora of molecular interactions resulting in a complex and heterogeneous phenotype. The hallmark of CAD is the development and progression of atheromatous narrowing of the coronary artery with an increasing risk of plaque rupture, resulting in acute coronary occlusion. Current preventive therapy for individuals at risk is directed towards the management of their lipid profile, blood pressure and promoting a healthy lifestyle. Genome-wide association studies (GWAS) have rapidly expanded our knowledge and provided novel leads to gain insights into human biology, optimize risk management and devise new therapeutic strategies[2]. To date, 57 loci have been reported by genome-wide association studies for CAD, mainly driven by efforts of the CARDIoGRAM- and C4D-consortia[3]. These genetic associations have identified genes that are among the targets of known and possible novel CAD therapies such as LDLR and HMGCR (HMG-coA reducatase inhibitors, statins), PCSK9 (PCSK9 inhibitors) and IL6R (Tocilizumab)[4,5]. Genetic association analyses have also identified therapeutic targets for many other conditions as well (reviewed by Plenge et al.[4]).

To further build upon our biological knowledge of CAD, to facilitate the identification of additional therapeutic targets, and to gain novel insights in the causal relationships between other cardiovascular phenotypes, continuous efforts directed at expanding the number of genetic regions associated with CAD are of paramount importance. Therefore, we set 3 goals. 1) Validate and identify novel loci by follow-up of the top-signals identified in the previous GWAS by the CARDIoGRAM-C4D consortium 2) determine biological pathways and candidate genes underlying the genome wide associated loci and 3) evaluate the association of the variants with common risk factors of CAD and common cardiovascular disorders to gain more insight into potential mediators of CAD per locus and trait.

RESULTS

First, we identified UK Biobank individuals with and without CAD. The prevalence and incidence of CAD conditions and events was captured by data collected at the Assessment Centre in-patient Health Episode Statistics (HES) and at any of the visits. A detailed definition of CAD can be found in the methods section and supplementary material.

Naturally, non-CAD individuals defined the control population but to improve statistical power we excluded individuals from the control population if their mother, father or sibling was reported to suffer from 'heart disease'. We validated this approach by constructing a genetic risk score based on the 57 previously reported loci weighted

with the effect estimates of the CARDIoGRAMplusC4D 1000 Genomes analysis assuming an additive model. The genetic risk score was associated with a family history of heart disease (n_{cases} = 71,263, $n_{controls}$ = 76,535, $p = 3\times10^{-128}$) in UK Biobank. Moreover, increased significance was observed for the association between the genetic risk score and CAD after excluding participants in the control group based on a family history ($p = 5\times10^{-183}$), compared to including these individuals ($p = 2\times10^{-147}$). Indicating that incorporating family history into the phenotype definition increases statistical power to detect associations between genetic variants and CAD.

This approach identified a total of 10,898 CAD cases and 76,535 non-CAD controls in UK Biobank that were imputed to the 1000 Genomes and UK10K reference panel[6]. The average age for CAD identified participants was 61.5 years and 55.8 for the controls. Detailed baseline characteristics are presented in Supplementary Table 1. To account for potential population stratification and genotyping differences, all associations in this manuscript were adjusted for the first 15 principle components, genotyping chip, gender and age.

Replication and identification of novel CAD loci

To date, 57 loci have been associated with CAD[3]. We performed logistic regressions between CAD status and these 57 previously reported CAD loci: 42 loci replicated at FDR < 0.05 in UK Biobank (Supplementary Table 2). A schematic overview of the 2-stage design to identify new CAD loci is presented in Figure 1. We first clumped genetic variants on LD (r^2<0.05, 1000 Genomes phase 1 v3 European panel) that reached a P value of < 1×10^{-5} in the latest CARDIoGRAMplusC4D GWAS. This resulted in 161 independent sets of variants sharing 120 independent loci (defined as 1MB at either side of the sentinel genetic variant; Supplementary Table 3). Seventy-one genetic variants in 52 loci were

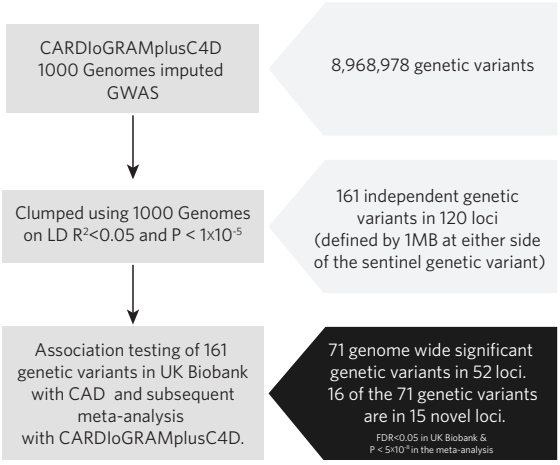


Figure 1. Flowchart of the study design

significantly associated (FDR < 0.05) with CAD in UK Biobank, were directionally concordant with CARDIoGRAMplusC4D, and were genome wide significant ($p < 5\times10^{-8}$) in the joint (meta-) analysis of UK Biobank and CARDIoGRAMplusC4D. Of these 52 loci, 37 were previously reported as genome-wide significant loci for CAD leaving 15 novel genome wide significant loci (Table 1, Supplementary Figure 1). All of the 15 novel loci were common. The minor allele frequency was above 7.6% with relatively weak effect sizes.

Table 1. Fifteen novel genome wide associated loci for coronary artery disease. Data presented here is from the meta-analysis; full summary statistics are available in Supplementary Table 3

Region	Genetic variant	EA / NEA	EAF	OR (95% CI)	P Value	Gene
1q21.3	rs11810571	G/C	0.787	1.069(1.05-1.09)	1.72×10^{-10}	TDRKH[n,e]
3p21.31	rs7623687	A/C	0.859	1.069(1.04-1.09)	3.28×10^{-08}	RHOA[n],AMT[n],TCTA[n],CDHR4[c],KLHDC8B[d]
3q21.2	rs142695226	G/T	0.136	1.078(1.05-1.10)	1.70×10^{-10}	UMPS[n,e],ITGB5[n,d]
3q25.2	rs433903	G/A	0.857	1.081(1.06-1.11)	6.06×10^{-10}	SGEF(Arhgef26)[n],DHX36[e]
4q21.21	rs10857147	T/A	0.269	1.061(1.04-1.08)	4.29×10^{-10}	PRDM8[n],FGF5[n]
4q27	rs11723436	G/A	0.305	1.053(1.04-1.07)	7.01×10^{-09}	MAD2L1[n],PDE5A[e]
4q31.21	rs35879803	C/A	0.702	1.051(1.03-1.07)	3.83×10^{-08}	ZNF827[n,e]
6p22.3	rs35541991	C/CA	0.312	1.049(1.03-1.07)	2.57×10^{-08}	HDGFL1[n]
11p15.2	rs1351525	T/A	0.674	1.049(1.03-1.07)	4.09×10^{-08}	ARNTL[n,c,e]
12q13.13	rs11170820	G/C	0.076	1.098(1.06-1.13)	4.09×10^{-08}	HOXC4[n]
12q24.31	rs2244608	G/A	0.349	1.056(1.04-1.07)	1.86×10^{-10}	HNF1A[nc],OASL[d]
14q24.3	rs3832966	I/D	0.458	1.054(1.04-1.07)	5.80×10^{-10}	TMED10[n,e],ZC2HC1C[e],RPS6KL1[e],NEK9[e],EIF2B2[e],ACYP1[e]
16q23.1	rs33928862	D/I	0.506	1.049(1.03-1.07)	2.47×10^{-08}	BCAR1[n,e,d]
16q23.3	rs7500448	A/G	0.772	1.069(1.05-1.09)	4.83×10^{-11}	CDH13[n,e,d]
19q13.2	rs138120077	D/I	0.140	1.072(1.05-1.10)	9.44×10^{-09}	HNRNPUL1[n,e],TGFB1[e,d],CCDC97[e]
19q13.2	rs8108632*	T/A	0.484	1.052(1.03-1.07)	9.54×10^{-09}	TGFB1[n,d],B9D2[n]

Abbreviations: EA = effect allele, NEA = Non-effect allele,, *EAF = effect allele frequency*, OR = Odds Ratio, *CI = confidence interval*, *I = Indel, D = Deletion. Candidate gene superscripts indicate the method of identification (n* = nearest gene, c=coding gene, d = depict gene, e = eQTL gene). * denotes the secondary signal in locus of region 19q13.2.

Candidate genes and pathway analyses

We prioritized 104 candidate genes in the 52 loci: 70 genes were prioritized based on proximity (the nearest gene and any additional gene within 10kb), 13 genes by coding genetic variants in linkage disequilibrium (r^2>0.8) with the sentinel genetic variant (Supplementary Table 4) and 51 genes based on expression quantitative trait loci (eQTL) analyses (Supplementary Table 5). Finally, 25 candidate genes were prioritized based on DEPICT analyses[7] (Supplementary Table 6). The DEPICT framework also identified 458 reconstituted gene sets that can be captured in 48 meta-genesets (Figure 2), the most significant gene set was 'abnormal vitelline vasculature morphology' (Supplementary Table 7);

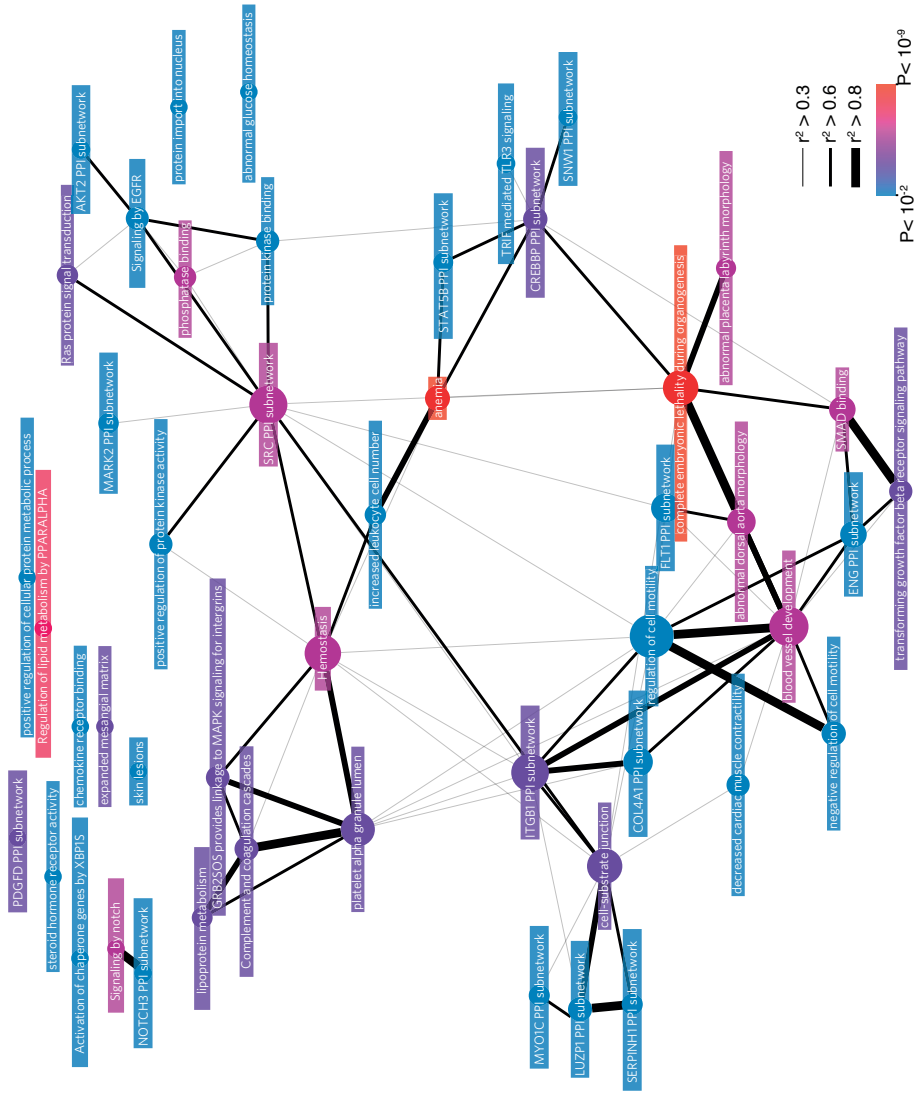


Figure 2. Gene-networks of the meta gene-sets that DEPICT prioritized at FDR < 0.05
Sizes of the nodes reflect the eigenvector centrality, an indicator of a node's centrality in the network.

'Arteries' was the only significantly prioritized tissue-type (FDR < 0.01, Supplementary Table 8). Network analysis identified 'regulation of cell motility' as the meta-geneset that was most central among all meta-genesets, together with 'negative regulation of cell motility' and 'blood vessel development', suggesting these pathways play an important role in CAD. More specific processes identified by DEPICT included hemostasis, anemia and increased leukocyte cell number. The function of each novel candidate gene has been summarized in the Supplementary Note.

Table 2. Associations in UK Biobank (N=143,936) between the genetic risk score based on the 71 genome wide significant CAD variants and cardiovascular profile

Phenotype	N individuals (%)	Beta (linear Regression) or odds ratio (logistic regression) (95% CI)	P value
Body-mass index	143,442 (99.7%)	-0.08(-0.14 to -0.02)	5.80×10^{-03}
Height	143,595 (99.8%)	-0.16(-0.24 to -0.09)	2.78×10^{-05}
Resting heart rate	135,946 (94.4%)	-0.45(-0.59 to -0.31)	2.03×10^{-10}
Blood pressure			
Systolic	143,770 (99.9%)	0.61(0.40 to 0.82)	1.27×10^{-08}
Diastolic	143,770 (99.9%)	-0.02(-0.14 to 0.10)	7.60×10^{-01}
Pulse pressure	143,770 (99.9%)	0.63(0.48 to 0.78)	1.65×10^{-16}
Mean arterial pressure	143,770 (99.9%)	0.19(0.05 to 0.33)	7.14×10^{-03}
Arterial stiffness index	54,184 (37.6%)	-0.02(-0.08 to 0.04)	5.38×10^{-01}
Smoking current	18,282 (14.5%)	-0.05(0.91 to 0.98)	5.45×10^{-03}
Medical Conditions			
Coronary Artery Disease	10,898 (8.2%)	2.21(2.11 to 2.32)	1.76×10^{-237}
Hypertension	48,927 (51.5%)	1.18(1.15 to 1.22)	5.89×10^{-35}
Diabetes	10,486 (7.9%)	1.10(1.05 to 1.16)	5.91×10^{-05}
Myocardial Infarction	5,145 (3.7%)	2.35(2.20 to 2.51)	2.05×10^{-143}
Heart failure	2,143 (1.5%)	1.43(1.29 to 1.58)	2.76×10^{-12}
Cardiomyopathy	522 (0.4%)	1.02(0.84 to 1.25)	8.39×10^{-01}
Atrial fibrillation / flutter	5,279 (3.8%)	1.10(1.03 to 1.18)	2.90×10^{-03}
Cerebral Infarction and TIA	4,043 (2.9%)	1.16(1.08 to 1.25)	4.68×10^{-05}
Device implantation	1,606 (1.1%)	1.51(1.35 to 1.69)	1.76×10^{-12}
Medication			
Beta-blockers	10,576 (7.9%)	1.42(1.36 to 1.49)	1.15×10^{-49}
Calcium channel-blockers	10,993 (8.3%)	1.17(1.12 to 1.23)	3.19×10^{-11}

Effect estimates with 95% Confidence Interval (CI) are shown as odds ratios for categorical variables (current smoking, cardiovascular disease, atherosclerosis, hypercholesterolemia, hypertension, diabetes, myocardial infarction, heart failure, atrial fibrillation / flutter, cerebral Infarction and TIA, device implantation, beta-blockers and calcium-channel blockers) and β estimates for quantitative variables (body-mass index, resting heart rate, systolic and diastolic blood pressure, pulse pressure, mean arterial pressure and arterial stiffness index). Abbreviation: N=Number, CI=Confidence Interval, TIA=Transient Ischemic Attack.

Mediating effects of CAD variants

CAD is a complex multifactorial disease, sharing biology with other atherosclerotic manifestations and vascular diseases. Therefore, we examined the association between genetic risk for CAD and several common cardiovascular phenotypes. We constructed a weighted genetic risk score by summing the number of CAD increasing risk alleles after multiplying the alleles with the corresponding β (based on the CARDIoGRAMplusC4D GWAS[3]). The genetic risk score that was based on the 71 independent genetic variants was associated with multiple other cardiovascular phenotypes and risk factors in UK Biobank (Table 2). At baseline, the genetic risk score was significantly associated with BMI, height, systolic blood pressure, mean arterial pressure, pulse pressure and heart rate. The genetic risk score was also associated with the presence of heart failure, hypertension, smoking, device implantation, cerebral infarction (including transient ischemic attack), atrial fibrillation and diabetes; whereas cardiomyopathy, diastolic blood pressure and arterial stiffness index were not significantly associated. We also tested whether the genetic risk score could predict cardiovascular mortality, death from coronary artery disease, death from myocardial infarction, and death from heart failure. The genetic risk score significantly predicted these outcomes in cox's proportional hazards models (Table 3). In addition, the genetic risk score predicted recurrent CAD events (n = 3,733) in participants with a history of CAD (n = 6,440; HR=1.10, Confidence interval 1.03 – 1.19, p = 0.009). Results were the same for a genetic risk score based on 57 previously known loci (Supplementary Table 9).

Table 3. Cox survival model predicting hazard of death using the genetic risk score based on the 71 genome wide significant CAD variants

Phenotype	N deaths (%)	Hazard Ratio (95% CI)	P value
Coronary Artery Disease	723 (0.5%)	1.75(1.48 to 2.08)	4.90×10^{-11}
Myocardial Infarction	210 (0.1%)	1.93(1.41 to 2.63)	3.40×10^{-05}
Heart failure	219 (0.2%)	1.39(1.02 to 1.88)	3.69×10^{-02}
Cardiomyopathy	40 (0.0%)	1.18(0.58 to 2.40)	6.48×10^{-01}
Cerebral Infarction and TIA	124 (0.1%)	1.32(0.87 to 1.98)	1.89×10^{-01}
All cause mortality	4373 (3.0%)	1.02(0.95 to 1.09)	5.65×10^{-01}
Cardiovascular mortality (as primary cause of death)	892 (0.6%)	1.46(1.26 to 1.70)	9.08×10^{-07}

Abbreviations: N=Number, CI=Confidence Interval, TIA=Transient Ischemic Attack

BOLT-REML[8] was used to assess the cumulative contribution of common genetic variants to CAD risk, and to estimate the degree of genetic correlation between CAD and other cardiovascular phenotypes in UK Biobank. We estimated the heritability of CAD by genome wide genetic variants, h^2_g , to be 0.063 (SE = 0.0046), which is 0.22 on the liability scale (with a prevalence of 0.076, based on UK Biobank). CAD and almost

all other studied cardiovascular phenotypes were genetically correlated, which led to comparable conclusions as our genetic risk score analyses. Narrow sense heritability estimates of all studied cardiovascular phenotypes and estimates of shared genetic correlations with CAD are available in Table 4.

Table 4. Heritability estimates for cardiovascular traits and the shared heritability between each trait and CAD in all UK Biobank Participants based on common genetic variation under the additive model (h^2_g). For dichotomous traits, the heritability on the observed 0-1 scale was transformed to h^2_g on the unobserved continuous liability scale by a linear transformation

Phenotype	$h2_g$(Se)	P-$h2_g$	Genetic correlation with CAD (Se)	P-correlation
Body-mass index	0.316 (0.005)	0	0.289 (0.026)	3.67×10^{-28}
Height	0.614 (0.004)	0	-0.142 (0.019)	5.93×10^{-13}
Resting heart rate	0.204 (0.005)	0	0.099 (0.033)	4.64×10^{-03}
Blood pressure		0		
Systolic	0.198 (0.005)	0	0.380 (0.034)	8.04×10^{-29}
Diastolic	0.197 (0.005)	0	0.316 (0.034)	3.15×10^{-20}
Pulse pressure	0.220 (0.005)	0	0.265 (0.030)	8.08×10^{-18}
Mean arterial pressure	0.191 (0.005)	0	0.380 (0.035)	3.70×10^{-27}
Arterial stiffness index	0.082 (0.012)	2.13×10^{-11}	0.075 (0.118)	3.26×10^{-01}
Smoking current	0.238 (0.012)	3.06×10^{-87}	0.258 (0.043)	8.30×10^{-09}
Medical Conditions				
Coronary Artery Disease	0.216 (0.016)	8.77×10^{-42}	-	-
Hypertension	0.310 (0.008)	0	0.577 (0.031)	1.39×10^{-75}
Diabetes	0.345 (0.017)	2.90×10^{-95}	0.412 (0.041)	2.88×10^{-23}
Myocardial Infarction	0.190 (0.025)	4.21×10^{-14}	1.000 (0.035)	2.40×10^{-174}
Heart failure	0.098 (0.044)	3.49×10^{-02}	0.679 (0.158)	4.04×10^{-05}
Cardiomyopathy	0.071 (0.134)	3.47×10^{-01}	0.462 (0.483)	2.53×10^{-01}
Atrial fibrillation / flutter	0.238 (0.025)	1.92×10^{-21}	0.323 (0.058)	8.29×10^{-08}
Cerebral Infarction and TIA	0.090 (0.028)	2.62×10^{-03}	0.635 (0.134)	4.93×10^{-06}
Device implantation	0.074 (0.055)	1.61×10^{-01}	0.678 (0.261)	1.36×10^{-02}
Medication				
Beta-blockers	0.156 (0.016)	1.94×10^{-22}	0.818 (0.053)	6.26×10^{-53}
Calcium channel-blockers	0.251 (0.016)	6.41×10^{-55}	0.547 (0.049)	2.76×10^{-28}

Abbreviations: $h2_g$= heritability based on genome wide variation, Se=Standard error, TIA=Transient Ischemic Attack. (h2g); standard error (SE); N.A.; not applicable.

To gain further insights into potential mediating mechanisms at the genetic variant level, we queried the GWAS-catalog for previously reported genome wide associations: for the novel loci, genetic variants in linkage disequilibrium (r^2>0.5) with rs10857147 (*FGF5*) was previously associated with blood pressure and serum urate levels; rs2244608

(*HNF1A/OASL*) with a wide range of biomarkers including lipids and urate levels; rs3832966 (*TMED10/NEK9*) with adult stature; rs1351525 with menarche; rs33928862 with pulmonary function; and rs8108632 (*B9D2/TGFB1/AXL*) with migraine and colorectal cancer risk (Supplementary Table 10). Furthermore, we performed association testing in the CAD controls of UK Biobank (N = 76,535) to reduce potential reverse causation of CAD for the following traits: systolic- and diastolic blood pressure, mean arterial pressure, pulse pressure, arterial stiffness index, heart rate, smoking and diabetes. We also performed lookups in previously published datasets of large GWAS: lipids[9], BMI[10], Hip circumference[11], waist-hip ratio[11] (adjusted for BMI), results are presented in Supplementary Table 11. Of the 71 genetic variants, 63 were nominally associated ($p < 0.05$) with one or more phenotypes. These lookups confirmed our findings of the GWAS-catalog query; rs2244608 is highly associated with total cholesterol and LDL ($p = 9\times10^{-21}$), rs10857147 (*FGF5*) is associated with blood pressure ($p = 2\times10^{-13}$) but also identified novel associations for the novel loci such as rs7500448 (pulse pressure, $p = 4\times10^{-11}$), rs1351525 (Systolic blood pressure $p = 7\times10^{-7}$) and rs33928862 (Systolic blood pressure $p = 5\times10^{-4}$). It also highlighted 21 of 71 variants without any association ($P>0.05$) with blood pressure or lipid traits (Supplementary Table 11).

DISCUSSION

Using a 2-stage design, adding 10,898 new cases and 76,535 controls to the 60,801/130,681 controls/cases previously studied by the CARDIoGRAMplusC4D consortium, we identified 15 novel loci reaching genome wide significance[3]. The variants of these 15 loci were common, with generally low effect sizes. In keeping with previous observations, our strategy did not reveal CAD variants of low frequency (MAF < 1-0.05%), suggesting that even other reference sets, techniques or larger sample sizes are required[3]. We added a relatively modest increase in cases (17.9%) compared to CARDIoGRAMplusC4D data but the number of additional controls was substantially higher (58.6%) and by filtering on a family history of 'heart disease' we might have decreased the number of misclassifications. These aspects of our strategy may have contributed to the relative large number of novel CAD loci compared to the latest CARDIoGRAMplusC4D that identified 10 new loci. Within UK Biobank we observed that the genetic risk score significantly predicts - and has a shared heritability with - a range of cardiovascular phenotypes, illustrating for example that genetically predicted CAD also increases risk for heart failure and atrial fibrillation, in line with observations from clinical practice.

Of the novel prioritized candidate genes, some have been previously reported for their involvement in blood vessel development. For example, *RHOA*, part of the Ras protein family, is involved in a multitude of cellular processes via the Rho-kinase pathway which

has a primary role in the regulation of contraction in vascular smooth muscle cells and promoting development of vascular remodeling[12]. *CDH13* which encodes T-cadherin, is a regulator of vascular wall remodeling, angiogenesis and is essential for adiponectin's vascular actions[13]. *TGFB1*, one of the most widely studied genes, is crucial in embryonic development and tissue homeostasis. The role of *TGFB1* in angiogenesis is a fact and long thought to play a role in CAD development, but the exact molecular pathways are hard to tackle due to the complex and multifactorial nature[14]. Rs2241718 near *TGBF1* has been prioritized previously as a functional regulatory variant[15] but is in low linkage disequilibrium (r2 <0.05) with the two signals identified here. Identifying two independent variants in this locus provides new opportunities to study the role of *TGFB1* in CAD. The product of *ITGB5*, integrin β5 has been studied in some detail for its role in cell adhesion and integrin-mediated signaling. It is believed that *ITGB5* is able to exert pro-angiogenic effects by enhancing the binding capacity of circulating angiogenic cells to endothelial cells[16]. Our pathway analyses also suggest that factors related to angiogenesis ('blood vessel development' and 'regulation of cell motility') are indeed central among the CAD loci, supplementing previously performed pathway analyses[17].

We also identified a novel CAD locus (rs433903) harboring *SGEF*. *SGEF* has been described to contribute to the formation of ICAM-1-induced endothelial docking structures that facilitate transendothelial migration and adhesion of leukocytes[18]. This process has an unfavorable role in atherosclerosis: *SGEF*$^{-/-}$ mice demonstrate a significant reduction in the formation of atherosclerotic plaque and was suggested as a novel therapeutic target, also since there appeared to be no other negative phenotypes [18,19]. Here, we demonstrate that rs433903 near *SGEF* is associated with CAD in humans and is not convincingly associated with other phenotypes such as blood pressure and lipids. Future studies are necessary to determine the exact molecular mechanisms underlying rs433903 and whether this variant is causally implicated in CAD through mechanisms of *SGEF* to further establish *SGEF* as a new candidate target for therapy.

The majority of preventive CAD medication is currently directed towards lowering LDL cholesterol and blood pressure, both of which are also closely associated with CAD on a genetic level, and considered to be causally related[20–22]. Genetic variants lacking any association with blood pressure or lipids might be of increased interest to be considered as novel (first in class) therapeutic targets that act independently from blood pressure or lipid lowering medication. However, our analyses are limited by the associative nature. To establish further evidence of the true causal genes and mechanisms underlying each association, further functional experiments are essential.

We are the first to have observed a significant genetic correlation between CAD and heart failure. The degree of shared heritability between CAD and heart failure was estimated to be as high as 0.68. We also observed that genetic risk for CAD was strongly associated with the occurrence of heart failure due to CAD, and predicts death of heart

failure with similar effects. It is well known that CAD plays a major role in heart failure, prevention of CAD is essential to maintaining functional myocyte reserve and preventing left ventricular systolic dysfunction[23]. Furthermore, a significant correlation and shared heritability was observed between the genetic risk score of CAD and increased risk of atrial fibrillation, perhaps due to atrial infarction but shared mechanisms of inflammation may also be responsible[24].

We could not only explain death due to CAD using our genetic risk score, in line with other studies[25], but could even predict progression of CAD as indicated by the significant association with recurrent CAD. A genetic risk score may be helpful to discriminate individuals at high risk of CAD and to direct more intensive preventive therapies. Future studies should be focused at replicating the newly identified loci and at further elucidating the molecular and pathophysiological mechanisms underlying CAD.

In summary, we report 15 novel loci, representing a 20% expansion of loci that are genome wide associated with CAD, including 2 independent variants near *TGFB1*. We also highlight widespread sharing of genetic variation between CAD and numerous other common cardiovascular diseases including atrial fibrillation and heart failure.

MATERIALS AND METHODS

UK biobank individuals

UK Biobank recruited participants with an age range of 40-69 years that registered with a general practitioner of the UK National Health Service (NHS). Between 2006–2010, in total 503,325 individuals were included. All study participants provided informed consent and the study was approved by the North West Multi-centre Research Ethics Committee. Detailed methods used by UK Biobank have been described elsewhere.

Ascertainment of resting coronary artery disease and controls

The prevalence and incidence of coronary artery disease conditions and events were captured by data collected at the Assessment Centre in-patient Health Episode Statistics (HES). CAD was defined using the following ICD 10 codes: I21-I25 covering ischaemic heart diseases and the following Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4 (OPCS-4) codes: K40-K46, K49, K50 and K75 which includes replacement, transluminal balloon angioplasty, and other therapeutic transluminal operations on coronary artery and percutaneous transluminal balloon angioplasty and insertion of stent into coronary artery. The exact phenotype definitions of UK Biobank are described in the Supplementary Note under section "Definitions used for UK Biobank analyses". Individuals from the control group were excluded if their

mother, father or sibling were reported to suffer from 'heart disease' to increase the true CAD/non-CAD ratio for our analysis.

Genotyping and imputation

Of the 500 thousand individuals with phenotype data in UK Biobank, 152,249 (25%) are currently genotyped. Genotyping, quality control and imputation was performed by UK Biobank and described in detail elsewhere[6,26]. Briefly, genotyping of 102,326 individuals was performed using the UK Biobank Axiom array (Affymetrix), and an additional 49,923 individuals were genotyped as part of the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) project. The Welcome Trust Centre for Human Genetics performed quality control before imputation and imputed the dataset using a merged reference panel of 1000 Genomes Phase 3 and UK10K[6]. The imputed dataset consisted of 72,355,667 genetic variants. For this work, genetic variants were included only if the imputation quality was greater than 0.3 and MAF>0.005 in line with the CARDIoGRAMplusC4D analysis, leaving 12,248,858 genetic variants. Participants were excluded based on gender mismatch, high missingness and high heterozygosity (n=662). We also removed 8,874 individuals based on relatedness (3rd degree or closer[6]), one of each related pair was excluded based on the highest missingness.

Statistical analysis

We selected genetic variants for replication from the CARDIoGRAMplusC4D[3] GWAS (downloaded from: http://www.cardiogramplusc4d.org/downloads) by filtering on $p < 1\times10^{-5}$ and linkage disequilibrium using the PLINK clumping procedure (*'--clump-kb 5000 --clump-r2 0.05'*, 1000 Genomes phase 1), after which we determined the number of 2-Megabase-loci, by assigning 1 Megabase regions at either side of the highest associated variant per locus (designated the sentinel genetic variant). Logistic regression analyses between genetic variants and the 10,898 CAD cases and 76,535 controls in UK Biobank were performed after adjustments for age, sex, the first 15 Principal Components to control for population stratification, and the genotyping array used. To account for multiple testing and declare novel loci we applied a replication p of FDR < 0.05 in UK Biobank and a genome wide significance threshold of $p < 5\times10^{-8}$ in the inverse-variance meta-analysis between the summary statistics of UK Biobank and CARDIoGRAMplusC4D.

Pathway analyses

The DEPICT Framework was used to identify enriched pathways, prioritize candidate genes at each loci and selects relevant tissues/cell types from co-expression networks of genes underlying the associated loci[7] (see Pers et al.[7] for a detailed description). We applied DEPICT on CARDIoGRAMplusC4D results at $p < 1\times10^{-5}$ which identified 194 independent loci using default settings (PLINK parameters, *'--clump-p1 1e-5 --clump-kb*

500 --clump-r2 0.01'), containing 489 genes. The gene prioritization, gene set enrichment and tissue/cell type enrichment analyses were run using the default settings in DEPICT (1000G dataset). We applied the affinity propagation method[27] to identify correlated genesets and for each correlated group the exemplar-geneset, which was named 'meta-geneset', and used Gephi (www.gephi.org)[28] to visualize the pearson correlation between pathways and calculate the centrality measures of each node (Figure 2).

Genetic risk score analyses & (shared) heritability of CAD

To study the relationship of CAD with other cardiovascular phenotypes, we created a weighted genetic risk score by summing the number of CAD risk-increasing alleles weighted (multiplied) for its β (estimated using the 1000 genomes meta-analysis[3]) of each associated genetic variant; assuming an additive effect. We performed a linear or logistic regression adjusted for age, gender, principle components and genotyping chip between the genetic risk score and cardiovascular phenotype. Cox regression analysis adjusted for age, gender, principle components and the genotyping chip was used to evaluate the predictive power of the genetic risk score on mortality and recurrent CAD events. Bivariate REML analyses were performed using BOLT-REML[8] to estimate the heritability of CAD and the genetic correlation of CAD with other cardiovascular traits. All directly genotyped variants that passed quality control were extracted from the imputed dataset (to ensure 100% call rate) and pruned on linkage disequilibrium ($r^2 < 0.05$) to obtain roughly 500k variants, as recommended[8]. Liability scale was estimated for dichotomous traits using linear transformation[29]. Gender, age, principle components and genotyping chip we included as covariates in all analyses.

Identification of candidate genes

We prioritized candidate genes for each of the 71-genome wide significant variants that were shared in 52 loci, based on the following criteria:

(1) The nearest gene or any gene located within 10kb of the sentinel genetic variant
(2) Any gene containing protein coding variants in linkage disequilibrium ($r^2 > 0.8$, UK Biobank) with the sentinel genetic variant (Supplementary Table 4).
(3) Expression QTL (eQTL) analyses in cis; we search for eQTLs (sentinel genetic variants or genetic variants in linkage disequilibrium, $r^2 > 0.8$, UK Biobank) in an eQTL dataset that was compiled from multiple tissues, including those of GTEX v6[30], STARNET[31] and large eQTL datasets of blood[32–34] (see Supplementary Table 5). We only considered eQTLs for which the top-eQTL was in linkage disequilibrium ($r^2 > 0.8$, UK Biobank) with the sentinel genetic variant and for which the eQTL $p < 1\times10^{-6}$.
(4) DEPICT-genes (see section "pathway analyses" for more details and Supplementary Table 6-8).

REFERENCES

1. Task Force Members *et al.* 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur. Heart J.* 34, 2949–3003 (2013).
2. Kullo, I. J. *et al.* Incorporating a Genetic Risk Score into Coronary Heart Disease Risk Estimates: Effect on LDL Cholesterol Levels (the MIGENES Clinical Trial). *Circulation* CIRCULATIONAHA.115.020109 (2016). doi:10.1161/CIRCULATIONAHA.115.020109
3. Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* 47, 1121–1130 (2015).
4. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through human genetics. *Nat. Rev. Drug Discov.* 12, 581–594 (2013).
5. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet Lond. Engl.* 379, 1214–1224 (2012).
6. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource. (2015).
7. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* 6, 5890 (2015).
8. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* 47, 284–290 (2015).
9. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* 45, 1274–1283 (2013).
10. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–206 (2015).
11. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 518, 187–196 (2015).
12. Shimokawa, H., Sunamura, S. & Satoh, K. RhoA/Rho-Kinase in the Cardiovascular System. *Circ. Res.* 118, 352–366 (2016).
13. Parker-Duffen, J. L. *et al.* T-cadherin Is Essential for Adiponectin-mediated Revascularization. *J. Biol. Chem.* 288, 24886–24897 (2013).
14. Zeng, L., Dang, T. A. & Schunkert, H. Genetics links between transforming growth factor β pathway and coronary disease. *Atherosclerosis* 253, 237–246 (2016).
15. Miller, C. L. *et al.* Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. *Nat. Commun.* 7, 12092 (2016).
16. Leifheit-Nestler, M. *et al.* Overexpression of Integrin β5 Enhances the Paracrine Properties of Circulating Angiogenic Cells via Src Kinase–Mediated Activation of STAT3. *Arterioscler. Thromb. Vasc. Biol.* 30, 1398–1406 (2010).
17. Ghosh, S. *et al.* Systems Genetics Analysis of GWAS reveals Novel Associations between Key Biological Processes and Coronary Artery Disease. *Arterioscler. Thromb. Vasc. Biol.* 35, 1712–1722 (2015).
18. Samson, T. *et al.* The Guanine-Nucleotide Exchange Factor SGEF Plays a Crucial Role in the Formation of Atherosclerosis. *PLoS ONE* 8, (2013).
19. Bitoun, P. Sgef controls macular, corpus callosum and hippocampal function and development, liver homeostasis, functions of the immune system, fever response atherosclerosis and tumorogenic cell growth. (2012).

20. Studies, T. I. C. for B. P. G.-W. A. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478, 103–109 (2011).
21. Do, R. *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat. Genet.* 45, 1345–1352 (2013).
22. Lieb, W. *et al.* Genetic predisposition to higher blood pressure increases coronary artery disease risk. *Hypertension* 61, (2013).
23. Doshi, D. *et al.* Underutilization of Coronary Artery Disease Testing Among Patients Hospitalized With New-Onset Heart Failure. *J. Am. Coll. Cardiol.* 68, 450–458 (2016).
24. Hu, Y.-F., Chen, Y.-J., Lin, Y.-J. & Chen, S.-A. Inflammation and the pathogenesis of atrial fibrillation. *Nat. Rev. Cardiol.* 12, 230–243 (2015).
25. Khera, A. V. *et al.* Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *N. Engl. J. Med.* 0, null (2016).
26. Eppinga, R. N. *et al.* Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. *Nat. Genet.* 48, 1557–1563 (2016).
27. Bodenhofer, U., Kothmeier, A. & Hochreiter, S. APCluster: an R package for affinity propagation clustering. *Bioinformatics* 27, 2463–2464 (2011).
28. Bastian, M., Heymann, S. & Jacomy, M. *Gephi: An Open Source Software for Exploring and Manipulating Networks*. (2009).
29. Lee, S. H., Wray, N. R., Goddard, M. E. & Visscher, P. M. Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* 88, 294–305 (2011).
30. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45, 580–585 (2013).
31. Franzén, O. *et al.* Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* 353, 827–830 (2016).
32. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 45, 1238–1243 (2013).
33. Bonder, M. J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *bioRxiv* (2015). doi:10.1101/033084
34. Zhernakova, D. *et al.* Hypothesis-free identification of modulators of genetic risk factors. *bioRxiv* (2015). doi:10.1101/033217

SUPPLEMENTARY APPENDIX

Supplementary Note 84

Summary of the candidate genes in the novel associated loci 84

Definitions used for UK Biobank analyses 87

Supplementary Figures 90

Supplementary References 91

SUPPLEMENTARY NOTE

Summary of the candidate genes in the novel associated loci

Locus 1q21.3, rs11810571 (TDRKH): The function of TDRKH is unknown in the literature, whereas the functions of the isoforms encoded by RORC, RORγ and RORγt are complex and widely studied [1].

Locus 3p21.31, rs7623687 (RHOA, AMT, TCTA, CDHR4 and KLHDC8B): RHOA (Ras homolog gene family, member A) is a small GTPase protein from the Rho family. The effects of RHOA are not all known it is primarily associated with cytoskeleton regulation, mostly actin stress fibers formation and actomyosin contractility. The RhoA/Rho-associated coiled-coil-forming kinase (ROCK) pathway participates in acute myocardial infarction and inhibiting of this pathway with atorvastatin improves the post-infarct microenvironment[2]. The AMT gene provides instructions for making an enzyme called aminomethyltransferase. This is one of four subunits that make up glycine cleavage enzyme. This complex is active in mitochondria. Mutations in this gene are responsible for ~15% of the glycine encephalopathy[3]. Only a few studies have been studying the role of TCTA (T-cell leukemia translocation-altered), it has been reported to play a role in human tumorigenesis and osteoclastogenesis[4] and inhibit proliferation of fibroblast-like synoviocytes[5]. CDHR4 (cadherin-related family member 4) cadherins are calcium-dependent cell adhesion proteins. They preferentially interact with themselves in a homophilic manner in connecting cells; cadherins may contribute to the sorting of heterogeneous cell types. KLHDC8B encodes a protein which forms a distinct beta-propeller protein structure of kelch domains (allowing for protein-protein interactions). Mutations have been associated with Hodgkin lymphoma.

Locus 3q21.2, rs142695226 (ITGB5 and UMPS): The product of ITGB5, integrin β5 is widely studied for it's role in cell adhesion and integrin-mediated signaling. It plays a role in angiogenesis, overexpression promotes new blood vessel formation in vivo by enhancing the binding capacity of circulating angiogenic cells to endothelial cells, among other molecular effects [6]. UMPS encodes a uridine 5'-monophosphate synthase, it catalyzes the reaction of orotic acid and ribose-5-phosphate to uridine monophosphate (UMP), an energy-carrying molecule.

Locus 3q25.2, rs433903 (ARHGEF26(SGEF) and DHX36): ARHGEF26 (Rho Guanine Nucleotide Exchange Factor 26) encodes a member of the Rho-guanine nucleotide exchange factor (Rho-GEF) family. These proteins regulate Rho GTPases by catalyzing the exchange of GDP for GTP; ARGHEF26 is also named SGEF, reported to play a crucial role in atherosclerosis and is suggested to be a potential therapeutic target[7]. DHX36 is a gene which is a member of the DEAH-box family of RNA-dependent NTPases. It may be involved in regulation of telomere length, function in sex development and spermatogenesis and may play a role in ossification [genecards].

Locus 4q21.21, rs10857147 (PRDM8 and FGF5): PRDM8 encodes a protein that belongs to a conserved family of histone methyltransferases that acts predominantly as negative regulators of transcription [genecards]. FGF5 is a member of the fibroblast growth factor family that play an important role in cell proliferation and differentiation; FGF5's major role is in regulation of hair length[8].

Locus 4q27, rs11723436 (MAD2L1 and PDE5A): MAD2L1 is a component of the mitotic spindle assembly checkpoint it prevents the anaphase, until all chromosomes at the metaphase are aligned. PDE5A (Phosphodiesterase 5A) is a Protein Coding gene. It is involved in the regulation of intracellular concentrations of cyclic nucleotides and is important for smooth muscle relaxation in the cardiovascular system. PDE5 expression is increased in patients with advanced cardiomyopathy[9].

Locus 4q31.22, rs35879803 (ZNF827): ZNF827 is a largely unknown zinc finger protein. It has been reported to recruit the NuRD (Nucleosome Remodeling Deacetylase) complex that has chromatin remodeling and histone deacetylase activities. The NuRD-ZNF827 complex promotes telomere-telomere recombination, it integrates and controls multiple mechanistic elements of 'alternative lengthening of telomeres' (ALT) activity [10].

Locus 6p22.3, rs35541991 (HDGFL1): HDGFL1 encodes hepatoma-derived growth factor-like 1. Variants near HDGFL1 have been genome wide associated with Total iron binding capacity[11] but its function remains to be determined.

Locus 11p15.2, rs1351525 (ARNTL): ARNTL (Aryl Hydrocarbon Receptor Nuclear Translocator Like), a transcriptional activator, and its product BMAL1, form the core components of the circadian clock and mainly known for interactions with CLOCK genes.

Locus 12q13.13, rs11170820 (HOXC4): There is not much known about HOXC4. HOXC4, is one of several HOXC genes located in a cluster on chromosome 12; three genes, HOXC5, HOXC4 and HOXC6, share a 5' non-coding exon. The homeobox genes encode a highly conserved family of transcription factors that play an important role in morphogenesis in all multicellular organisms. [genecards]. HOXC4 has been studied in relationship to differentiation of hematopoietic stem cells[12] and adipose tissue[13].

Locus 12q24.31, rs2244608 (HNF1A, OASL): HNF1A is a frequent cause of monogenic diabetes (MODY-HNF1A) and highly expressed in liver, pancreas and the proximal tubule of the kidney. It has been shown to be highly associated with lipid levels[14], and suggested to be involved in CRP, GGT, and other atherosclerotic and metabolic risk factors[15]. It plays a major role in the expression of various hepatic, renal, and pancreatic genes/proteins including megalin (Low density lipoprotein-related protein 2), cubilin [16], PCSK9[17]. Altogether, HNF1A is a pleiotropic gene that is widely studied with many functions. OASL encods oligoadenylate synthetase enzymes, which are cytoplasmic dsRNA sensors belonging to the antiviral innate immune system.

Locus 14q24.3, rs3832966 (TMED10, NEK9, ZC2HC1C, RPS6KL1, EIF2B2 and ACYP1): Little is known about TMED10's function. TMED10 is thought to be a type I membrane

protein that is localized to the plasma membrane and golgi cisternae, involved in vesicular protein trafficking. NEK9 recently reported to be a cause for a lethal skeletal dysplasia. Loss of function results in defects of fibroblasts including a reduced proliferation capability and delayed cell cycle progression through the G1/S boundary and S-phase and could also be involved in ciliopathy[18]. EIF2B2 (Eukaryotic Translation Initiation Factor 2B Subunit Beta) is a Protein Coding gene. Diseases associated with EIF2B2 include Leukoencephalopathy With Vanishing White Matter and Late Infantile Cach Syndrome. Among its related pathways are Gene Expression and Translation Insulin regulation of translation [genecards]. ACYP1 is a member of the acylphosphatase family. The encoded protein is a small cytosolic enzyme that catalyzes the hydrolysis of the carboxyl-phosphate bond of acylphosphates. Two isoenzymes have been isolated and described based on their tissue localization: erythrocyte (common) type acylphosphatase encoded by this gene, and muscle type acylphosphatase [genecards]. nothing is known about the function of ZC2HC1C (Zinc Finger C2HC-Type Containing 1C) and RPS6KL1 (Ribosomal Protein S6 Kinase Like 1).

Locus 16q23.1, rs33928862 (BCAR1): Breast cancer anti-estrogen resistance protein 1 is a protein that in humans is encoded by the BCAR1 gene and involved in various cellular events, basic signaling of developmental/physiological processes and involved in regulation homeostasis of various tissues, BCAR1's functions and role has been reviewed previously[19]. A variant in LD (r^2=0.65), rs4888378, has been associated with Carotid Intima-Media Thickness and coronary artery disease risk[20] .

Locus 16q23.3 rs7500448 (CDH13): CDH13 is a widely studied member of de cadherin family, it is an adhesion glycoprotein known as T-cadherin and is recognized as an LDL receptor, although different to other LDL recepters, it activates Erk 1/2 tyrosine kinase and the nuclear translocation of NF-kappaB[2122]. GVs near this gene have previously been genome wide associated with blood pressure[23] and adiponectin levels[24] ($P=6.8\times10^{-165}$), among others. The locus has also been identified in one of the first genome wide association studies of coronary artery disease[25], although not at genome wide significance. None of the reported SNPs were in LD (r^2>0.001) with the current finding, rs7500448. We identified rs7500448 to be highly associated ($P=8\times10^{-13}$) with pulse pressure in UK Biobank.

Locus 19q13.2, rs138120077 and rs8108632 (B9D2, TGFB1, HNRNPUL1 and CCDC97): not much is known about B9D2's function, the encoded protein localizes to basal bodies and cilia, mutations cause Meckel syndrome [26]. TGFB1, transforming growth factor beta1, is one of the most widely studied genes. It is a multifunctional peptide which regulates proliferation, differentiation, adhesion, migration, among other functions and studied for its role in angiogenesis, cardiovascular syndromes and vascular biology [27–29]. rs2241718 near TGBF1 has been prioritized as a functional regulatory variant[30] but is in low LD with the 2 signals identified in our study. The heterogeneous nuclear ribonucleo-

protein U-like 1 (HNRPUL1) gene encoding for a hetero-geneous ribonuclear protein believed to be involved in mRNA processing and transport [31,32], candidate studies found significant associations between variants and CAD in high risk people with familial hypercholesteromelia [33]. Nothing is known for CCDC97 (Coiled-Coil Domain Containing 97), but it has been recently studied as a candidate for regulatory mechanisms of CAD, together with TGFB1[30]. This study showed that while the 3'-untranslated region variant at CCDC97/TGFB1, rs2241718, was predicted to affect binding, this variant might not alter endogenous CCDC97 levels, but rather serve as an enhancer for neighboring TGFB1 in human coronary artery smooth muscle cells.

Definitions used for UK Biobank analyses

Prevalent and incident coronary artery disease (CAD), hypercholesterolemia, hypertension, diabetes, myocardial infarction (MI), heart failure, atrial fibrillation / flutter, cerebral infarction was derived from self-reported (touchscreen questionnaire and verbal interview) and/or the diagnosis was captured using the Hospital Episode Statistics (HES) records (using the following ICD codes: I21 - I25 for CAD; E78 for hypercholesterolemia; I10 - I15 for hypertension; E10 - E14 for diabetes; I21 - I22 for MI; I42, I150 for heart failure; I48 for atrial fibrillation and flutter; I63 - I64 for cerebral infraction and transient ischemic attack; all codes beginning with I as primary diagnosis were used to define cardiovascular mortality. In addition information for CAD and device implantation (pacemaker or implantable cardioverter defibrillator) was captured through HES records which are coded according to the Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4 (OPCS-4) (using the following OPCS-4 codes: K40 - K46, K49, K50, K75 for CAD and; K59 - K61 and U31 for device implantation. The "Spell and Episode" category contains data relating to diagnoses made during hospital in-patient stay. It includes main and secondary diagnoses, coded according to the International Classification of Diseases (ICD). The main diagnosis corresponds to be the main reason for the hospital admission, while secondary diagnoses are more often contributory or underlying conditions. We used both the main and secondary diagnoses for recording prevalent and incident risk factors, conditions and events. For defining the control group we excluded participants who reported that their mother, father or sibling suffered form 'heart disease' (Field ID 20107, 20110 and 20111). Information on smoking status was collected using the touchscreen questionnaire at baseline visit. Medication usage was collected at the baseline visit during a verbal interview by a trained nurse on prescription medications (Field ID 20003). Data on beta block-blocker and calcium channel-blocker therapy was defined with corresponding medication codes (beta-blockers and calcium channel blockers, please see below for the exact codes that were used). Body mass index was calculated using BMI value constructed from height and weight measured during the initial Assessment Centre visit (Field ID 21001) and Body

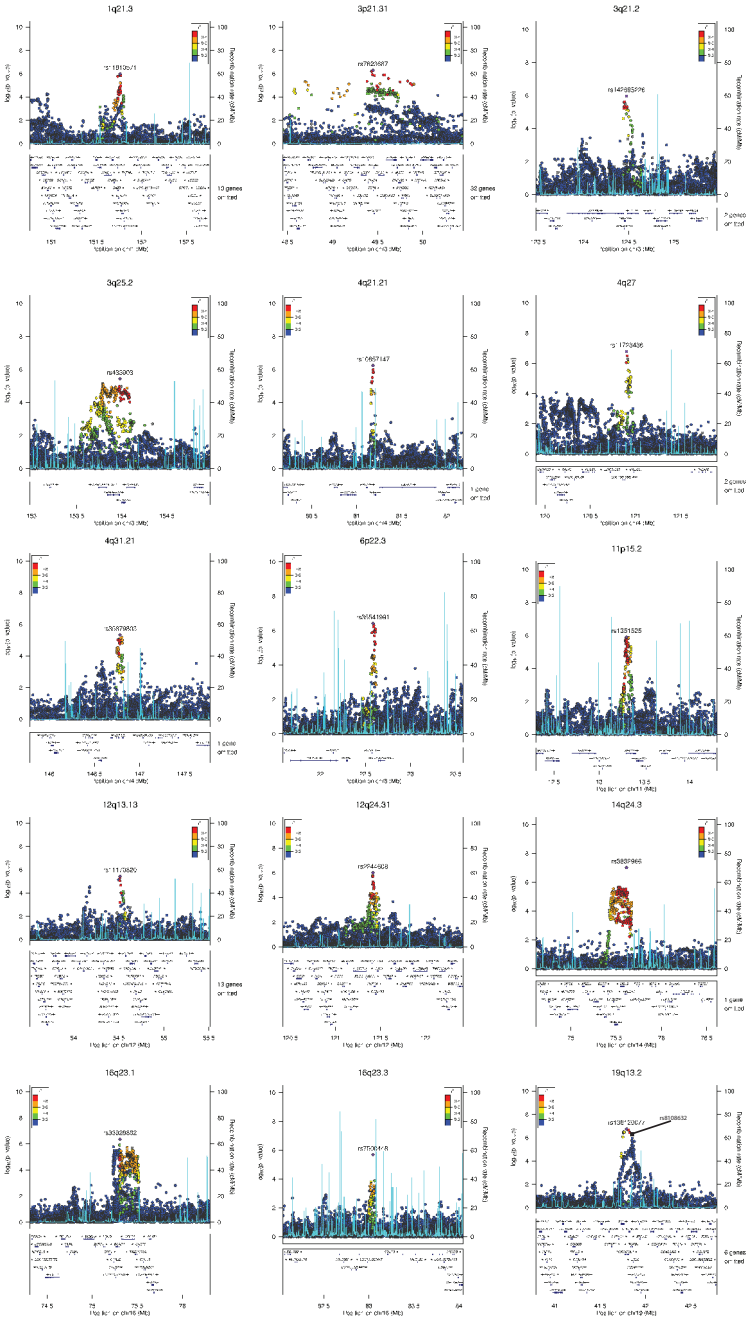
composition estimation by impedance measurement (Field ID 23104). Blood pressure was measured using the manual reading (Field ID 93, 94) and automated reading (Field ID 4079, 4080) measurements. Pulse pressure was calculated by subtracting the diastolic from the systolic blood pressure value. Mean arterial pressure (MAP) was calculated by MAP = [(2 x diastolic blood pressure) + systolic blood pressure] divided by 3. When multiple measurements during firs visit were available the mean of all measurements were averaged and used in the analysis. In the UK Biobank cohort PWV for ASI assessment was measured using the PulseTrace PCA2 (CareFusion, San Diego, USA) (Field-ID 21021). The PulseTrace PCA2 uses finger photoplethysmography to obtain the pulse waveform during a 10-15 seconds measurement using an infrared sensor clipped to the end of the index finger[34]. When multiple measurements were available the mean of all measurements were averaged and used in the analysis.

Beta-blocker medication codes: 1140866724, 1140866738, 1140860192, 1140860292, 1140860404, 1140860308, 1140860312, 1141194804, 1140860316, 1140860320, 1140860322, 1140860332, 1140860336, 1140860340, 1141194808, 1140860342, 1140860418, 1140860422, 1140860426, 1140864950, 1140909368, 1141164276, 1141162898, 1141169516, 1141184722, 1140879758, 1140879760, 1140879762, 1140879818, 1140879822, 1140879824, 1140879830, 1140879834, 1140879842, 1140879854, 1140879866, 1141180778, 1141146124, 1141146126, 1141194810, 1141146128, 1140866692, 1140916342, 1140866704, 1140866764, 1140866766, 1140851556, 1140866778, 1140866782, 1140866784, 1140866798, 1140866802, 1140866800, 1140866804, 1140916730, 1140916868, 1140917076, 1141152076, 1140866712, 1141156754, 1141156808, 1141172742, 1140866726, 1140866756, 1140860172, 1140864410, 1140922930, 1140860232, 1140860244, 1140860250, 1140860266, 1140860274, 1140860278, 1140860180, 1140860194, 1140860212, 1140851576, 1140851480, 1140860220, 1140860222, 1140851484, 1140860230, 1140910614, 1140860294, 1140851492, 1140860304, 1140860362, 1140860380, 1140860382, 1140860386, 1140860390, 1140860394, 1140860396, 1140860398, 1140860400, 1140860402, 1140860406, 1140860410, 1140860314, 1140860318, 1140851508, 1140860324, 1140860328, 1140860330, 1140860334, 1140860338, 1140916628, 1140860348, 1141146184, 1140860352, 1140860356, 1140860358, 1140860434, 1140860492, 1141171152, 1141184324, 1141182904, 1141187780, 1140851522, 1140863724, 1140860498, 1141168498, 1141164280, 1141187048

Calcium-channel-blocker medication codes: 1141165470, 1141150926, 1141153328, 1140926778, 1140851784, 1140861088, 1140861114, 1140911088, 1141150538, 1141157140, 1141169730, 1140861190, 1140879802, 1140888646, 1140861276, 1140928226, 1141153394, 1140872568, 1140879806, 1140879810, 1140888510,

1141153026, 1140861128, 1140851730, 1140861130, 1140861136, 1140861138, 1140861166, 1140926780, 1141157136, 1140911698, 1141151474, 1140917428, 1140917452, 1141153454, 1140923618, 1140861176, 1140861090, 1140923572, 1140851794, 1140926188, 1140926966, 1140861110, 1140927934, 1140927940, 1140861120, 1141145870, 1141150500, 1140916930, 1141152600, 1141166752, 1141162546, 1140851798, 1140851800, 1140861194, 1140861202, 1141200400, 1140928212, 1141187094, 1141188152, 1141188576, 1141188836, 1141188920, 1141190160, 1141199858, 1141200782, 1141201814, 1140861282, 1140928234, 1141153032, 1141153400, 1141167832, 1141175224, 1141171804, 1141174684, 1141180238, 1141173766, 1141187962, 1141188936, 1141190548

Supplemental tables (excel file) are available online.



Supplementary Figure 1. Regional plots of the 15 novel genome wide associated loci with CAD
LD (R^2) was based on the Europeans of 1000 Genomes Phase 1 v3. P-values were based on the CARDIoGRAMplusC4D GWAS data to provide an accurate overview of the P-value distribution among variants at each locus.

SUPPLEMENTARY REFERENCES

1. Cook, D. N., Kang, H. S. & Jetten, A. M. Retinoic Acid-Related Orphan Receptors (RORs): Regulatory Functions in Immunity, Development, Circadian Rhythm, and Metabolism. *Nucl. Recept. Res.* 2, (2015).
2. Zhang, Q. *et al.* Atorvastatin treatment improves the effects of mesenchymal stem cell transplantation on acute myocardial infarction: the role of the RhoA/ROCK/ERK pathway. *Int. J. Cardiol.* 176, 670–679 (2014).
3. Van Hove, J., Coughlin, C. & Scharer, G. in *GeneReviews(®)* (eds. Pagon, R. A. et al.) (University of Washington, Seattle, 1993).
4. Kotake, S., Yago, T., Kawamoto, M. & Nanke, Y. The role of T-cell leukemia translocation-associated gene protein in human tumorigenesis and osteoclastogenesis. *J. Biomed. Biotechnol.* 2012, 675317 (2012).
5. Nanke, Y. *et al.* A novel peptide from TCTA protein inhibits proliferation of fibroblast-like synoviocytes of rheumatoid arthritis patients. *Cent.-Eur. J. Immunol.* 39, 468–470 (2014).
6. Leifheit-Nestler, M. *et al.* Overexpression of Integrin β5 Enhances the Paracrine Properties of Circulating Angiogenic Cells via Src Kinase–Mediated Activation of STAT3. *Arterioscler. Thromb. Vasc. Biol.* 30, 1398–1406 (2010).
7. Samson, T. *et al.* The Guanine-Nucleotide Exchange Factor SGEF Plays a Crucial Role in the Formation of Atherosclerosis. *PLoS ONE* 8, (2013).
8. Higgins, C. A. *et al.* FGF5 is a crucial regulator of hair length in humans. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10648–10653 (2014).
9. Pokreisz, P. *et al.* Ventricular phosphodiesterase-5 expression is increased in patients with advanced heart failure and contributes to adverse ventricular remodeling after myocardial infarction in mice. *Circulation* 119, 408–416 (2009).
10. Conomos, D., Reddel, R. R. & Pickett, H. A. NuRD–ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination. *Nat. Struct. Mol. Biol.* 21, 760–770 (2014).
11. Li, J. *et al.* Genome-wide admixture and association study of serum iron, ferritin, transferrin saturation and total iron binding capacity in African Americans. *Hum. Mol. Genet.* 24, 572–581 (2015).
12. Xin, C., Zhao, C., Yin, X., Wu, S. & Su, Z. Bioinformatics analysis of molecular mechanism of the expansion of hematopoietic stem cell transduced by HOXB4/HOXC4. *Hematol. Amst. Neth.* 21, 462–469 (2016).
13. Singh, S., Rajput, Y. S., Barui, A. K., Sharma, R. & Datta, T. K. Fat accumulation in differentiated brown adipocytes is linked with expression of Hox genes. *Gene Expr. Patterns* 20, 99–105 (2016).
14. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* 45, 1274–1283 (2013).
15. Reiner, A. P. *et al.* Common coding variants of the HNF1A gene are associated with multiple cardiovascular risk phenotypes in community-based samples of younger and older European-American adults: the Coronary Artery Risk Development in Young Adults study and the Cardiovascular Health Study. *Circ. Cardiovasc. Genet.* 2, 244–254 (2009).
16. Terryn, S. *et al.* Tubular proteinuria in patients with HNF1α mutations: HNF1α drives endocytosis in the proximal tubule. *Kidney Int.* 89, 1075–1089 (2016).
17. Shende, V. R. *et al.* Reduction of circulating PCSK9 and LDL-C levels by liver-specific knockdown of HNF1α in normolipidemic mice. *J. Lipid Res.* 56, 801–809 (2015).
18. Casey, J. P. *et al.* Recessive NEK9 mutation causes a lethal skeletal dysplasia with evidence of cell cycle and ciliary defects. *Hum. Mol. Genet.* ddw054 (2016). doi:10.1093/hmg/ddw054

19. Camacho Leal, M. del P. *et al.* p130Cas/BCAR1 scaffold protein in tissue homeostasis and pathogenesis. *Gene* 562, 1–7 (2015).
20. Boardman-Pretty, F. *et al.* Functional Analysis of a Carotid Intima-Media Thickness Locus Implicates BCAR1 and Suggests a Causal Variant. *Circ. Cardiovasc. Genet.* 8, 696–706 (2015).
21. Rubina, K. A., Kalinina, N. I., Parfyonova, Y. V. & Tkachuk, V. A. T-cadherin as a receptor regulating angiogenesis and blood vessel remodeling. *Biochem. Mosc. Suppl. Ser. Membr. Cell Biol.* 1, 57–63 (2007).
22. Kipmen-Korgun, D. *et al.* T-cadherin mediates low-density lipoprotein-initiated cell proliferation via the Ca(2+)-tyrosine kinase-Erk1/2 pathway. *J. Cardiovasc. Pharmacol.* 45, 418–430 (2005).
23. Org, E. *et al.* Genome-wide scan identifies CDH13 as a novel susceptibility locus contributing to blood pressure determination in two European populations. *Hum. Mol. Genet.* 18, 2288–2296 (2009).
24. Wu, Y. *et al.* A meta-analysis of genome-wide association studies for adiponectin levels in East Asians identifies a novel locus near WDR11-FGFR2. *Hum. Mol. Genet.* 23, 1108–1119 (2014).
25. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678 (2007).
26. Dowdle, W. E. *et al.* Disruption of a ciliary B9 protein complex causes Meckel syndrome. *Am. J. Hum. Genet.* 89, 94–110 (2011).
27. ten Dijke, P. & Arthur, H. M. Extracellular control of TGFβ signalling in vascular development and disease. *Nat. Rev. Mol. Cell Biol.* 8, 857–869 (2007).
28. Goumans, M.-J., Liu, Z. & ten Dijke, P. TGF-β signaling in vascular biology and dysfunction. *Cell Res.* 19, 116–127 (2009).
29. Pardali, E., Goumans, M.-J. & Dijke, P. ten. Signaling by members of the TGF-β family in vascular morphogenesis and disease. *Trends Cell Biol.* 20, 556–567 (2010).
30. Miller, C. L. *et al.* Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. *Nat. Commun.* 7, 12092 (2016).
31. Gabler, S. *et al.* E1B 55-Kilodalton-Associated Protein: a Cellular Protein with RNA-Binding Activity Implicated in Nucleocytoplasmic Transport of Adenovirus and Cellular mRNAs. *J. Virol.* 72, 7960–7971 (1998).
32. Kzhyshkowska, J., Rusch, A., Wolf, H. & Dobner, T. Regulation of transcription by the heterogeneous nuclear ribonucleoprotein E1B-AP5 is mediated by complex formation with the novel bromodomain-containing protein BRD7. *Biochem. J.* 371, 385–393 (2003).
33. van der Net, J. B. *et al.* Replication study of 10 genetic polymorphisms associated with coronary heart disease in a specific high-risk population with familial hypercholesterolemia. *Eur. Heart J.* 29, 2195–2201 (2008).
34. UK Biobank. *UK Biobank Arterial Pulse-Wave Velocity*

PART II

METABOLOMICS

Chapter 5

Effect of Metformin Treatment on Lipoprotein Subfractions in Non-Diabetic Patients with Acute Myocardial Infarction: A Glycometabolic Intervention as Adjunct to Primary Coronary Intervention in ST Elevation Myocardial Infarction (GIPS-III) Trial

Ruben N. Eppinga, Minke H. Hartman, Dirk J. van Veldhuisen, Chris P. Lexis, Margery A. Connelly, Erik Lipsic, Iwan C.C. van der Horst, Pim van der Harst, Robin P. Dullaart

Adapted from PLoS One. 2016 Jan 25;11(1):e0145719

ABSTRACT

Objective: Metformin affects low density lipoprotein (LDL) and high density (HDL) subfractions in the context of impaired glucose tolerance, but its effects in the setting of acute myocardial infarction (MI) are unknown. We determined whether metformin administration affects lipoprotein subfractions 4 months after ST-segment elevation MI (STEMI). Second, we assessed associations of lipoprotein subfractions with left ventricular ejection fraction (LVEF) and infarct size 4 months after STEMI.

Methods: 371 participants without known diabetes participating in the GIPS-III trial, a placebo controlled, double-blind randomized trial studying the effect of metformin (500 mg bid) during 4 months after primary percutaneous coronary intervention for STEMI were included of whom 317 completed follow-up (clinicaltrial.gov Identifier: NCT01217307). Lipoprotein subfractions were measured using nuclear magnetic resonance spectroscopy at presentation, 24 hours and 4 months after STEMI. (Apo)lipoprotein measures were obtained during acute STEMI and 4 months post-STEMI. LVEF and infarct size were measured by cardiac magnetic resonance imaging.

Results: Metformin treatment slightly decreased LDL cholesterol levels (adjusted P=0.01), whereas apoB remained unchanged. Large LDL particles and LDL size were also decreased after metformin treatment (adjusted P<0.001). After adjustment for covariates, increased small HDL particles at 24 hours after STEMI predicted higher LVEF (P=0.005). In addition, increased medium-sized VLDL particles at the same time point predicted a smaller infarct size (P<0.001).

Conclusion: LDL cholesterol and large LDL particles were decreased during 4 months treatment with metformin started early after MI. Higher small HDL and medium VLDL particle concentrations are associated with favorable LVEF and infarct size.

INTRODUCTION

The clinical relevance of plasma lipids and lipoprotein levels in predicting (recurrent) coronary heart disease is well appreciated [1]. Indeed, pharmacological treatment aimed at lowering low density lipoprotein (LDL) cholesterol is an essential part of the routine care of patients with a history of myocardial infarction (MI) [2]. Importantly, lipoprotein particles are highly heterogeneous in size, structure and function with probable consequences for cardiovascular risk prediction [3–10]. In the non-acute setting, LDL and high density lipoprotein (HDL) particle characteristics have been proposed to be more closely associated with (incident) coronary heart disease compared to LDL cholesterol and HDL cholesterol concentrations [3–5,7–9,11–21]. When determined at presentation of MI, LDL cholesterol, HDL cholesterol and triglycerides have been variably shown to predict recurrent adverse cardiac events [10]. However, little is currently known about the prognostic value of lipoprotein subfraction characteristics obtained in the setting of an acute MI.

In subjects with impaired glucose tolerance, metformin administration modestly reduces the LDL particle concentration, and concomitantly decreases small dense LDL particles and increases small HDL particles, as determined by nuclear magnetic resonance (NMR) spectrometry [22]. Furthermore, metformin improves insulin resistance [23,24], which has been recently identified as a marker of adverse cardiac outcome [24,25]. Taken together these findings provide a rationale to determine whether metformin affects lipoprotein subfraction characteristics in patients with an acute coronary event.

The Glycometabolic Intervention in Adjunct to Primary Percutaneous Coronary Intervention in ST-Segment Elevation Myocardial Infarction (GIPS-III) trial was designed to evaluate the effect of 4 months metformin treatment on left ventricular function in non-diabetic patients with ST-segment elevation MI (STEMI) [26,27]. The rationale of this study is based on experimental findings showing that metformin may beneficially affect left ventricular function through activation of a number of intracellular pathways and alters mitochondrial function as outlined extensively elsewhere [26]. Among other potentially beneficial effects, metformin may also affect lipid and lipoprotein levels [26], which was a predetermined tertiary efficacy endpoint of the GIPS-III trial [27]. This randomized trial provides a framework to determine effects of metformin on lipoprotein metabolism, and to evaluate associations of lipoprotein subfractions, obtained in the setting of an acute MI, with left ventricular ejection fraction (LVEF) and infarct size assessed at 4 months.

The present anciliary analyses were initiated to test the extent to which metformin treatment affects lipoprotein subfraction characteristics in GIPS-III participants. Second, we determined the association of lipoprotein subfractions with LVEF and infarct size.

METHODS

Study population

The GIPS-III trial is has been registered as clinical trial with identifier: NCT01217307. The design and primary results of the GIPS-III trial has been reported in detail elsewhere [26,27]. The inclusion and exclusion criteria of the GIPS-III stuy are reported in Table 1. In brief, 380 non-diabetic patients undergoing primary percutaneous coronary intervention (PCI) for STEMI were randomized to receive a 4-month regimen with either metformin 500 mg twice daily or matching placebo twice daily. During the PCI procedure, all patients except one provided verbal informed consent followed by written informed consent. This subject was excluded, as well as were subjects in whom lipoprotein subfractions measurements were not available. As a result, 185 subjects receiving metformin and in 186 subjects receiving placebo were available for the current analyses. From these participants we determined lipoprotein subfractions in 371 subjects at baseline, 338 subjects 24 hours post-MI and 317 subjects 4 months post-MI. A total of 271 subjects completed 4 months follow-up evaluation by cardiac magnetic resonance imaging (MRI). From these subjects in 268 lipoprotein subfractions were determined at baseline, 250 subjects 24 hours post-MI and 257 4 subjects months post-MI.

Blood samples were obtained shortly after admission at the catheterization laboratory, after 24 hours post-MI and 4 months after randomization. Very low density lipoproteins (VLDL), LDL and HDL particle profiles were determined at these 3 time points. Samples for glucose, glycated hemoglobin (HbA1c), plasma total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, apolipoprotein (apo)B and apoA-I were obtained at the catheterization laboratory and 4 months after randomization.

Table 1. In- and exclusion criteria for the GIPS-III trial

Inclusion criteria	Exclusion criteria
• The diagnosis acute MI defined by chest pain suggestive for myocardial ischemia for at least 30 min, the time from onset of the symptoms less than 12 h before hospital admission, and an ECG recording with ST- segment elevation of more than 0.1 mV in 2 or more leads	• Prior MI • Diabetes • Creatinin >177 µ mol/L measured pre-PCI • Need for coronary artery bypass grafting
• Verbal followed by written informed consent	• Rescue PCI after thrombolytic therapy
• At least one stent sized ≥ 3.0 mm	• When subjects develop a condition which, in the investigator's judgment, precludes study therapy
• Eligible for cardiac MRI-scan: - Contra-indication to metformin - Body Mass Index <40 kg/m2 an estimated life-expectancy of less than 6 months - no ferromagnetic metal objects in the body - no claustrophobia	• Inability to provide informed consent • Younger than 18 years • Contra-indication to metformin • an estimated life-expectancy of less than 6 months

Laboratory measurements

Serum and EDTA-anticoagulated plasma samples were stored at – 80 °C until analyzed. Plasma total cholesterol, LDL cholesterol and HDL cholesterol were measured by a direct quantitative assay using cholesterol (PEG-) esterase and (PEG-) cholesterol oxidase on a Roche Modular *P* autoanalyzer (Roche Diagnostics, Indianapolis, IN, USA). Non-HDL cholesterol was calculated as the difference between total cholesterol and HDL cholesterol. Triglycerides (TG) were quantified using the LipoProfile-3 algorithm (LipoProfile-3 algorithm; LipoScience Inc. (now Labcorp Inc.), Raleigh, North Carolina, USA) [16]. Quantification of TG was accomplished by converting NMR particle numbers to lipid mass concentration units, assuming that the lipoprotein particles have normal lipid content. NMR-derived values correlate well with chemically measured values. Apolipoprotein (apo) B and apoA-I were computationally estimated by the use of the high-throughput 1H nuclear magnetic resonance (NMR) metabolomics platform of Computational Medicine (Oulu, Finland) [28].

VLDL, LDL and HDL particle profiles were measured by NMR spectroscopy with the LipoProfile-3 algorithm (LabCorp, Raleigh, North Carolina, USA), as described [16]. VLDL, LDL and HDL subclasses were quantified from the amplitudes of their spectroscopically distinct lipid methyl group NMR signals, and were expressed in concentration units, i.e. µmol/L or nmol/L. The lipoprotein subfraction particle concentrations are considered to represent an estimate of the respective lipoprotein particle numbers. Diameter range estimates were for VLDL: large VLDL (including chylomicrons if present; > 60 nm), medium VLDL (35 to 60 nm) and small VLDL (27 to 35 nm), for LDL: IDL (23 to 27 nm), large LDL (21.2 to 23 nm) and small LDL (18 to 21.2 nm), and for HDL: large HDL particles: 9.4 to 14 nm; medium HDL particles: 8.2 to 9.4 nm; small HDL particles: 7.3–8.2 nm. The VLDL, LDL and HDL particle concentrations were calculated as the sum of the respective lipoprotein subclasses. Weighted-average VLDL, LDL and HDL sizes were derived from the sum of the diameter of each subclass multiplied by its relative mass percentage based on the amplitude of its methyl NMR signal [16].

NT-proBNP was routinely measured with a sandwich immunoassay on a Roche Modular E platform (Mannheim, Germany).

Cardiac Magnetic Resonance Imaging (MRI)

LVEF and infarct size were measured by cardiac magnetic imaging [29]. These outcome measures were assessed by MRI four months after infarction. Details of the imaging analysis has been reported elsewhere [26,27]. An independent core laboratory (Image Analysis Center, Free University Medical Center, Amsterdam, The Netherlands) evaluated the MRI scans and assessed the primary efficacy measure, blinded for treatment allocation and clinical patient data.

Myocardial Blush Grade (MBG)

MBG was categorized as previously described [30]. A physician blinded to data analyzed coronary angiograms.

Statistical analysis

R (version 3.02 or higher, http://www.r-project.org/) was used for statistical analyses. Values for continuous variables that are normally distributed are presented as mean ± SD. Continuous variables not normally distributed are presented as median and interquartile ranges (IQRs).

Because not all lipoprotein subfractions were normally distributed (Shapiro-Wilk Normality test, $P < 0.05$), they were normalized using rank-based inverse normal transformation across all time points.

Pearson correlation coefficients were calculated from lipoprotein subfractions at 4 months after acute MI, and plotted using the corrplot function of the corrplot package of R. The correlation matrix is presented in Suppl. Fig. 1.

The extent to which clinical parameters, laboratory values and lipoprotein subfraction measurements between treatment (metformin and placebo) groups were significantly different at the various time-points (baseline, i.e. at admission for MI, 24 hours post-MI and 4 months post-MI) was determined after data normalization using unpaired T tests. Difference in medication use was assessed using a multinomial chi-squared test. None of the baseline lipoprotein variables were significantly different between the treatment groups.

For this reason the statistical comparisons of in the main results were given as the *P*-values of the unpaired T tests after 4 months of treatment in primary analysis. In addition, regression models were used to examine the changes in lipoprotein subfractions between the two treatment groups (placebo and metformin). In this analysis, the respective lipoprotein variable at 24 hours and after 4 months was the dependent variable with the following independent covariates: treatment assignment, age at randomization, sex, body mass index (BMI), statin use at 4 months and the baseline lipoprotein subfraction of interest.

For routine laboratory values the significance level was set at $P \leq 0.01$. In view of multiple testing of lipoprotein subfraction data, a principal components (PCs) analysis was carried out using the prcomp-function of R. The first 8 components explained 96% of the variation, of which the first 5 components explained 84% of the variation in the data set (Supplementary Data). On the basis of the PCs, the multiple testing corrected significance level of lipoprotein subfractions was set to $P \leq 0.05/8$ components, equivalent to $P \leq 0.0063$.

To examine the relationship between baseline laboratory values or lipoprotein subfraction levels at different time points (baseline and 24 hours post-MI) with LVEF or infarct size (4 months post-MI), linear regression models were used. In the primary analysis, LVEF or infarct size was the dependent variable with baseline laboratory values or lipoprotein subfractions as the independent variable. A secondary analysis was

performed adjusting for baseline laboratory values or lipoprotein subfractions, age at randomization, sex, baseline NT-proBNP concentration, treatment allocation, MBG and statin use at 4 months which we considered to be relevant covariates.

RESULTS

Clinical paramaters, MRI parameters and laboratory values

Table 2 summarizes clinical parameters, MRI parameters and laboratory measurements at baseline and at 4 months after intervention. There were no significant differences in clinical and laboratory characteristics between the two treatment groups at baseline. After 4 months of intervention there was a significantly lower LDL cholesterol in the metformin group (2.1 [1.8-2.4] mmol/L) group compared to the placebo group (2.2 [1.8-2.4 2.7] mmol/L); $P = 0.01$ after adjustment for baseline LDL cholesterol, age at randomization, sex, BMI, and statin use at 4 months). In a sensitivity analysis we performed a logistic regression with reduction of LDL cholesterol as dependent variable and metformin as independent variable; this also resulted in a trend towards lower LDL cholesterol ($\beta = -0.25$; SE: 0.12; $P = 0.04$).

Plasma triglycerides were much lower in both groups at baseline compared to the values after 4 months of follow-up ($P < 0.001$ for all comparisons), whereas total cholesterol, non-HDL cholesterol, LDL cholesterol and apoB were higher at baseline than after follow-up of 4 months ($P < 0.001$ for all comparisons).

Effect of metformin treatment on lipoprotein subfractions

Table 3 and Suppl. figure 2 show the median values of lipoprotein subfraction levels and lipoprotein sizes at baseline, 24 hours post-MI and 4 months post-MI. There were no significant differences between the two treatment groups at baseline and 24 hours post-MI. After 4 months of treatment, large LDL particles (270.5 [190.0-365.8] vs 170.0 [93.0-278] nmol/L and LDL size (20.3 [20.0-20.6] vs 20.5 [20.1-20.9] nm) were decreased in the metformin group compared to the placebo group ($P \leq 0.001$ for each). After correcting for their baseline values, age at randomization, sex, BMI and statin use at 4 months, these differences remained significant $P \leq 0.001$ for each). Figure 1 illustrates the changes in the various lipoprotein subfractions on lipoprotein particle size after 4 months of therapy with metformin compared to placebo. Remarkably, all VLDL subfractions and VLDL size were much lower in both groups at baseline compared to 24 hours and after 4 months of follow-up ($P < 0.001$ for all comparisons). Conversely, the LDL particle concentration, as well as the IDL and large LDL subfractions were higher at baseline than at after 4 months of follow-up ($P < 0.001$ for all comparisons). However, the HDL particle concentration and HDL subfractions were unaffected by metformin administration.

Table 2. Clinical-, MRI- and laboratory parameters by treatment group at baseline and after 4 months

	Baseline (Myocardial Infarction)				**4 Months Treatment**			
Clinical parameters	**Placebo (*n*=186)**	**Metformin (*n*=185)**	***P***		**Placebo (*n*=160)**	**Metformin (*n*=157)**	***P***	***P*****
Age, mean	58.72 ± 11.44	58.80 ± 11.82	0.95					
Women, n (%)	48 (26%)	46 (25%)	0.83					
Body Mass index, mean	27.0 ± 3.9	27.0 ± 3.8	0.92		27.4 ± 4.0	27.0 ± 4.0	0.37	
Statin use, n (%)	15 (8%)	13 (7%)	0.70					
Medication use	**at discharge**		0.25		**at 4 months**		0.92	
Statin, n (%)	185 (99%)	184 (99%)			150 (94%)	147 (94%)		
Aspirin, n (%)	181 (97%)	178 (96%)			154 (96%)	146 (93%)		
Thienopyridine, n (%)	186 (100%)	185 (100%)			157 (98%)	152 (97%)		
Coumarine , n (%)	7 (4%)	13 (7%)			7 (4%)	12 (8%)		
Beta-blocker, n (%)	181 (97%)	174 (94%)			150 (94%)	148 (94%)		
ACE-inhibitor or ARB, n (%)	141 (76%)	153 (83%)			139 (87%)	140 (89%)		
Calcium-channel blocker, n (%)	3 (2%)	9 (5%)			13 (8%)	11 (7%)		
Aldosterone antagonist, n (%)	12 (7%)	25 (14%)			19 (12%)	25 (16%)		
Laboratory values			***P***	***P****			***P***	***P*****
Glucose (mmol/L)	8.4 (7.0-9.8)	8.2 (7.0-9.4)	0.93	0.93	5.6 (5.2-6.1)	5.6 (5.1-6.2)	0.94	0.90
HbA1c (%)	5.8 (5.6-6.0)	5.8 (5.6-6.1)	0.51	0.43	5.9 (5.7-6.1)	5.9 (5.6-6.0)	0.11	0.07
Creatinine (µmol/L)	72 (62.3-80)	72 (61.0-84.0)	0.66	0.69	79.5 (72-88.3)	79 (70.0-87.0)	0.65	0.49
Total Cholesterol (mmol/L)	5.4 (4.8-6.0)	5.3 (4.7-6.1)	0.87	0.87	3.9 (3.5-4.4)	3.8 (3.4-4.2)	0.07	0.06
LDL Cholesterol (mmol/L)	3.8 (3.3-4.4)	3.7 (3.1-4.4)	0.57	0.65	**2.2 (1.8-2.7)**	**2.1 (1.8-2.4)**	**0.02**	**0.01**
non-HDL Cholesterol (mmol/L)	4.2 (3.7-4.9)	4.2 (3.6-4.9)	0.98	0.97	2.7 (2.7-3.2)	2.6 (2.2-3.0)	0.16	0.11
HDL Cholesterol (mmol/L)	1.1 (0.9-1.3)	1.1 (0.9-1.3)	0.94	0.97	1.2 (1.0-1.5)	1.1 (0.9-1.4)	0.14	0.09

Table 2. Clinical-, MRI- and laboratory parameters by treatment group at baseline and after 4 months (continued)

	Baseline (Myocardial Infarction)				4 Months Treatment			
Clinical parameters	**Placebo (*n*=186)**	**Metformin (*n*=185)**	***P***		**Placebo (*n*=160)**	**Metformin (*n*=157)**	***P***	***P*****
Triglycerides (mmol/L)	0.69 (0.57-0.88)	0.69 (0.59-0.98)	0.22	0.20	1.38 (1.04-1.84)	1.44 (1.02-2.01)	0.12	0.31
ApoB (g/L)	0.79 (0.68-0.93)	0.80 (0.64-0.93)	0.87	0.95	0.76 (0.66-0.86)	0.74 (0.66-0.87)	0.68	0.48
ApoA-I (g/L)	1.26 (1.16-1.40)	1.29 (1.16-1.42)	0.38	0.36	1.33 (1.23-1.47)	1.30 (1.21-1.44)	0.11	0.17
NT-proBNP (ng/L)	78 (37-175)	84 (42-235)	0.31	0.27	167 (74-355)	163 (67-389)	0.47	0.65
Cardiac MRI parameters								
LVEF, % (95% CI)	NA	NA			54.7 (53.4-56.1)	53 (51.5-54.6)	0.10	
LVEDV, ml (95% CI)	NA	NA			194.1 (186.4-201.9)	194.6 (186.5-202.6)	0.94	
LVESV, ml (95% CI)	NA	NA			89.8 (83.8-95.8)	93.6 (87.3-100.0)	0.39	
LVEDM, g (95% CI)	NA	NA			101.6 (97.2-106.0)	102.0 (98.1-105.8)	0.91	
Infarct size, % (95% CI)	NA	NA			8.8 (7.5-10.1)	8.8 (7.5-10.2)	0.97	

Abbreviations: Apo: apolipoprotein; ARB: aldosterone receptor blocker; BMI: body mass index; HbA1c: glycosylated hemoglobin; HDL: high density lipoproteins; LDL: low density lipoproteins; LVEF: left ventricular ejection fraction; LVEDV: left venticular end diastolic volume; LVESV: left ventricular end systolic volume; LVEDM: left ventricular end diastolic mass; non-HDL: non-high density lipoproteins; NT-proBNP: N-terminal pro brain natriuretic peptide. Baseline and 4 months data are presented as mean ± SD, median (interquartile range), and number (percentage) as appropriate. *P*-values for placebo vs. Metformin by unpaired T tests . *P*-value* adjusted for age at randomization, sex, BMI. *P*-value** additionally adjusted for respective baseline laboratory value and statin use at 4 months. Bold is *P*-value ≤ 0.01.

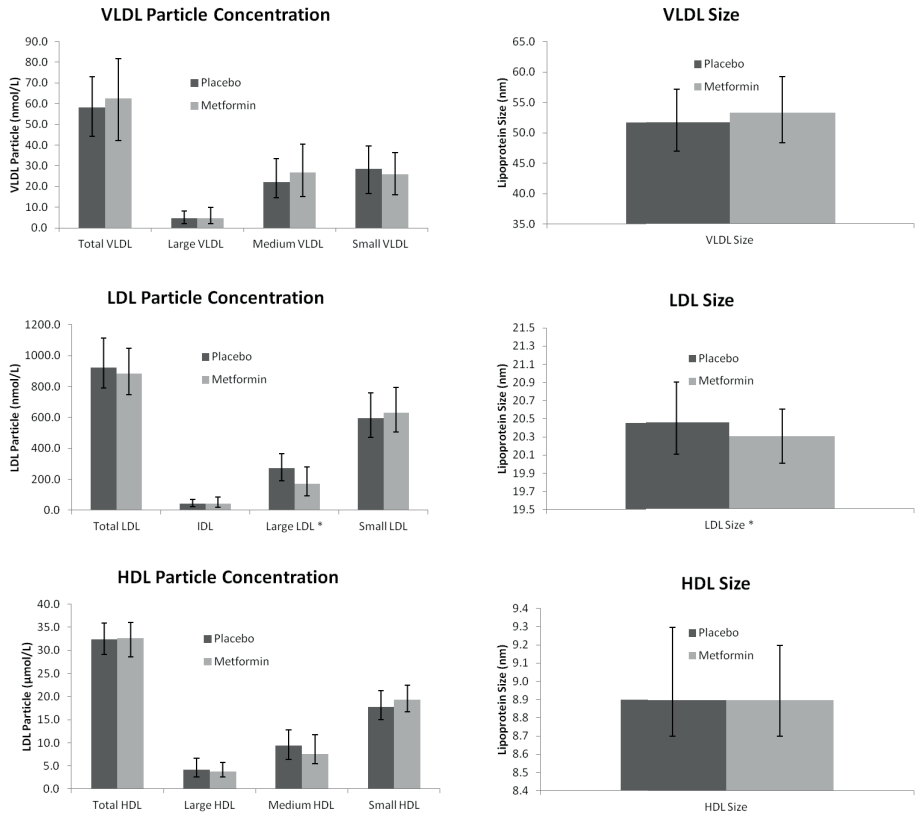


Figure 1. Lipoprotein subfraction concentrations and size using NMR after 4 months according to treatment group (VLDL-***P***; LDL-***P***; HDL-***P***; Lipoprotein size)
Data are presented as median (interquartile range). *P*-values from unpaired – tests. *P*-value ≤ 0.0063 placebo vs. Metformin group.

Relationship baseline laboratory values with left ventricular ejection fraction and infarct size

Table 4 shows the relationships of baseline laboratory values with LVEF and infarct size at 4 months post-MI. As reported previously [manuscript submitted] NT-proBNP at baseline was negatively associated with LVEF 4 months post-MI ($P = 0.008$). Glucose level at baseline was positively associated with infarct size ($P = 0.001$). After adjustment for age, sex, baseline NT-proBNP level, treatment allocation and myocardial blush grade this relationship remained significant ($P = 0.003$).

Table 3. Lipoprotein subfractions by treatment group at baseline, 24 hours and 4 months

	Baseline (Myocardial Infarction)				24 Hours post-MI				4 Months post-MI			
	Placebo (*n*=186)	Metformin (*n*=185)	*P*	*P**	Placebo (*n*=172)	Metformin (*n*=166)	*P*	*P***	Placebo (*n*=160)	Metformin (*n*=157)	*P*	*P****
VLDL particle concentration (nmol/L)	12.2 (3.5-36.5)	15.2 (5.5-42.0)	0.086	0.067	61.4 (45.9-82.7)	64.9 (48.6-84.7)	0.353	0.737	58.3 (44.1-72.9)	62.4 (42.2-81.7)	0.472	0.624
Large VLDL (nmol/L)	0.3 (0.2-0.9)	0.4 (0.2-1.6)	0.117	0.102	6.4 (3.3-10.5)	7.6 (3.9-11.8)	0.175	0.308	4.7 (2.1-8.2)	4.7 (2.1-10.0)	0.096	0.379
Medium VLDL (nmol/L)	0.7 (0.0-4.8)	1.7 (0.0-8.5)	0.242	0.220	27.3 (14.9-39.5)	28.0 (17.5-43.3)	0.642	0.902	22.2 (14.6-33.4)	26.7 (15.0-40.5)	0.112	0.186
Small VLDL (nmol/L)	10.4 (3.0-25.1)	11.3 (4.4-31.3)	0.271	0.247	24.1 (16.3-34.4)	25.4 (15.0-36.8)	0.579	0.594	28.4 (16.7-39.6)	26.0 (16.1-36.4)	0.121	0.047
VLDL Size (nm)	43.8 (37.8-48.1)	44.2 (39.9-50.6)	0.160	0.163	53.4 (47.4-59.4)	54.5 (48.3-60.6)	0.131	0.146	51.7 (46.9-57.2)	53.3 (48.3-59.2)	0.032	0.070
LDL particle concentration (nmol/L)	1509.0 (1292.8-1730.5)	1485.0 (1264.0-1756.0)	0.982	1.000	1262.0 (1055.5-1520.0)	1264.5 (1036.0-1504.5)	0.784	0.729	923.0 (790.5-1114.0)	883.0 (747.0-1048.0)	0.223	0.082
IDL (nmol/L)	102.5 (62.3-180.0)	89.0 (46.0-154.0)	0.080	0.080	81.5 (41.8-146.5)	76.0 (38.0-132.5)	0.295	0.691	40.5 (22.0-69.8)	43.0 (19.0-83.0)	0.713	0.659
Large LDL (nmol/L)	764.5 (546.8-975.0)	665.0 (481.0-898.0)	0.114	0.113	491.5 (330.0-653.3)	432.0 (268.3-597.8)	0.103	0.567	**270.5 (190.0-365.8)**	**170.0 (93.0-278.0)**	**0.000**	**0.000**
Small LDL (nmol/L)	606.5 (423.0-803.8)	659.0 (470.0-923.0)	0.074	0.058	675.0 (496.5-875.0)	708.5 (560.8-959.0)	0.145	0.540	596.5 (470.8-758.8)	631.0 (506.0-795.0)	0.065	0.352
LDL Size (nm)	21.1 (20.7-21.4)	21.0 (20.6-21.3)	0.118	0.112	20.8 (20.4-21.2)	20.7 (20.3-21.0)	0.064	0.297	**20.5 (20.1-20.9)**	**20.3 (20.0-20.6)**	**0.000**	**0.001**
HDL particle concentration (µmol/L)	28.8 (25.6-32.5)	29.4 (26.2-32.5)	0.113	0.111	28.0 (25.0-30.7)	28.0 (24.9-30.7)	0.873	0.201	32.4 (29.1-35.9)	32.6 (28.6-36.1)	0.688	0.403
Large HDL particles (µmol/L)	3.0 (2.1-4.6)	3.1 (2.1-4.9)	0.813	0.781	2.7 (1.7-4.9)	3.1 (1.9-4.7)	0.339	0.305	4.2 (2.6-6.6)	3.8 (2.6-5.8)	0.334	0.675
Medium HDL particles (µmol/L)	14.5 (9.6-20.0)	14.6 (10.1-19.7)	0.903	0.889	9.0 (6.5-11.7)	8.3 (6.0-11.0)	0.293	0.319	9.4 (6.4-12.8)	7.6 (5.5-11.7)	0.012	0.004
Small HDL particles (µmol/L)	9.5 (5.0-14.0)	10.8 (6.3-14.8)	0.266	0.273	15.1 (12.2-18.4)	16.1 (13.1-18.6)	0.559	0.667	17.8 (15.0-21.3)	19.3 (16.7-22.5)	0.022	0.027
HDL Size (nm)	9.0 (8.8-9.3)	9.0 (8.8-9.3)	0.952	0.936	8.7 (8.5-9.2)	8.8 (8.6-9.1)	0.334	0.257	8.9 (8.7-9.3)	8.9 (8.7-9.2)	0.438	0.847

Particle concentrations are presented for subgroups of each major lipid fraction. Size values present the average particle size across all lipoprotein subgroups. Baseline, 24 hours post-MI and 4 months post-MI data are presented as median (interquartile range). *P*-values for placebo vs. Metformin group by unpaired T tests. *P*-value* adjusted for age at randomization, sex, and body mass index. *P*-value** adjusted for baseline value of respective lipoprotein subfractions. *P*-value*** additionally adjusted for statin use at 4 months. Bold is *P*-value ≤ 0.0063.

Table 4. Relationship LVEF and infarct size with baseline laboratory values

	LVEF β (95% CI)	*P*	*P**	Infarct Size β (95% CI)	*P*	*P**
Glucose	-1.024 (-2.051, 0.002)	0.050	0.073	**1.577 (0.640, 2.514)**	**0.001**	**0.003**
HbA1c	0.551 (-0.513, 1.615)	0.309	0.172	-0.230 (-1.219, 0.760)	0.648	0.417
Total Cholesterol	-0.308 (-1.329, 0.712)	0.552	0.367	0.279 (-0.685, 1.243)	0.570	0.337
LDL Cholesterol	-0.288 (-1.306, 0.730)	0.578	0.278	0.198 (-0.762, 1.158)	0.685	0.379
non-HDL Cholesterol	-0.147 (-1.163, 0.869)	0.776	0.468	0.001 (-0.951, 0.953)	0.998	0.577
HDL Cholesterol	-0.385 (-1.424, 0.653)	0.466	0.840	0.806 (-0.170, 1.782)	0.105	0.272
Triglycerides	0.520 (-0.626, 1.667)	0.372	0.544	-0.509 (-1.589, 0.571)	0.354	0.730
ApoB	-0.381 (-1.417, 0.656)	0.470	0.475	-0.069 (-1.072, 0.935)	0.893	0.846
ApoA-I	-0.448 (-1.416, 0.521)	0.364	0.494	0.255 (-0.670, 1.180)	0.588	0.701
NT-proBNP	**-1.383 (-2.405, -0.362)**	**0.008**		0.745 (-0.217, 1.707)	0.128	

Linear regression model of LVEF or infarct size with baseline laboratory values. Unadjusted coefficients are shown. Abbreviations: Apo: apolipoprotein; HbA1c: glycosylated hemoglobin; HDL: high density lipoproteins; LDL: low density lipoproteins; LVEF: left ventricular ejection fraction; non-HDL: non-high density lipoproteins; NT-proBNP: N-terminal pro brain natriuretic peptide. *P*-value*: adjusted for age at randomization, sex, baseline NT-proBNP concentration, treatment allocation, myocardial blush grade and statin use at 4 months. Bold is *P*-value

Relationship of lipoprotein subfraction levels at baseline and 24 hours post-MI with left ventricular ejection fraction and infarct size

Table 5 shows the relationships of lipoprotein subfraction levels at baseline and 24 hours with LVEF and infarct size at 4 months after MI. None of the lipoprotein subfractions at baseline were associated with LVEF or infarct size. The concentration of small HDL particles 24 hours post-MI was positively associated with LVEF ($P = 0.003$) and negatively with infarct size ($P = 0.006$). Furthermore, the total VLDL particle concentration ($P = 0.003$) and the medium VLDL particle concentration ($P = 0.001$) 24 hours post-MI were both negatively associated with infarct size. After adjustment for age, sex, baseline NT-proBNP level, treatment allocation, myocardial blush grade and statin use, the relationship between small HDL particles and LVEF ($P = 0.005$), and between the medium VLDL particle concentration with infarct size ($P = 0.001$) remained significant. The association of the total VLDL particle concentration and of small HDL particles with infarct size was not significant after adjustment for these covariates.

DISCUSSION

We show here that 4 months metformin administration initiated directly after the acute phase of MI is associated with slight reductions reduced LDL cholesterol without affecting apoB levels in a large group of patients without previously established diabetes.

Table 5. Relationship of LVEF and infarct size at 4 months with lipoprotein subfraction levels at baseline and 24 hours.

	Baseline (Myocardial Infarction)						24 hours post Myocardial Infarction					
	LVEF β (95% CI)	*P*	*P**	Infarct Size β (95% CI)	*P*	*P**	LVEF β (95% CI)	*P*	*P**	Infarct Size β (95% CI)	*P*	*P**
VLDL particle concentration	0.843 (-0.213, 1.900)	0.117	0.152	-0.573 (-1.553, 0.407)	0.251	0.531	1.719 (0.333, 3.105)	0.015	0.015	**-1.875 (-3.107, -0.643)**	**0.003**	0.008
Large VLDL	0.257 (-0.942, 1.456)	0.673	0.601	-0.339 (-1.440, 0.762)	0.545	0.812	1.034 (-0.397, 2.466)	0.176	0.171	-1.278 (-2.567, 0.010)	0.052	0.086
Medium VLDL	0.852 (-0.384, 2.088)	0.176	0.179	-1.151 (-2.291, -0.012)	0.048	0.100	1.731 (0.374, 3.089)	0.012	0.007	**-2.186 (-3.407, -0.965)**	**0.001**	**0.001**
Small VLDL	0.667 (-0.229, 1.563)	0.144	0.190	-0.278 (-1.120, 0.565)	0.517	0.862	0.284 (-1.022, 1.590)	0.653	0.864	-0.089 (-1.276, 1.099)	0.883	0.820
VLDL Size	0.086 (-1.098, 1.270)	0.886	0.866	-0.294 (-1.379, 0.791)	0.594	0.701	0.550 (-0.741, 1.841)	0.399	0.394	-0.509 (-1.663, 0.644)	0.385	0.431
LDL particle concentration	-0.078 (-1.257, 1.101)	0.896	0.586	0.076 (-1.027, 1.178)	0.893	0.549	0.431 (-0.816, 1.677)	0.495	0.574	-0.314 (-1.438, 0.810)	0.583	0.799
IDL	-0.075 (-1.151, 1.001)	0.891	0.796	-0.742 (-1.741, 0.258)	0.145	0.153	0.595 (-0.456, 1.646)	0.265	0.487	-0.234 (-1.195, 0.727)	0.632	0.903
Large LDL	-0.996 (-2.063, 0.070)	0.067	0.028	1.069 (0.073, 2.065)	0.035	0.029	-0.438 (-1.720, 0.844)	0.500	0.212	1.076 (-0.097, 2.248)	0.072	0.041
Small LDL	0.762 (-0.138, 1.663)	0.097	0.101	-0.614 (-1.451, 0.223)	0.150	0.263	0.513 (-0.538, 1.563)	0.337	0.153	-0.885 (-1.825, 0.056)	0.065	0.064
LDL Size	-1.065 (-2.174, 0.045)	0.060	0.024	1.125 (0.094, 2.155)	0.033	0.034	-0.432 (-1.526, 0.662)	0.520	0.147	0.974 (-0.002, 1.949)	0.050	0.029
HDL particle concentration	0.185 (-0.896, 1.265)	0.737	0.462	-0.525 (-1.533, 0.483)	0.306	0.211	1.317 (0.110, 2.524)	0.031	0.051	-0.887 (-1.986, 0.211)	0.113	0.149
Large HDL particles	0.286 (-0.784, 1.357)	0.599	0.315	0.462 (-0.536, 1.460)	0.363	0.752	-0.264 (-1.301, 0.773)	0.632	0.661	0.335 (-0.599, 1.269)	0.480	0.740
Medium HDL particles	0.207 (-0.867, 1.280)	0.705	0.782	-0.542 (-1.542, 0.458)	0.287	0.377	-0.651 (-2.030, 0.727)	0.352	0.352	0.594 (-0.641, 1.829)	0.344	0.244
Small HDL particles	-0.264 (-1.382, 0.854)	0.642	0.888	0.127 (-0.922, 1.177)	0.811	0.985	**2.315 (0.815, 3.814)**	**0.003**	**0.005**	**-1.913 (-3.273, -0.554)**	**0.006**	0.009
HDL Size	0.414 (-0.853, 1.681)	0.520	0.275	0.503 (-0.686, 1.693)	0.405	0.780	-0.439 (-1.432, 0.555)	0.505	0.644	0.367 (-0.530, 1.264)	0.421	0.670

Linear regression model of left ventricular ejection fraction (LVEF) or infarct size with lipoprotein subfractions as Baseline (MI) and 24 hours post-MI. Unadjusted β coefficients are shown. *P*-value is from univariate linear regression model. *P*-value* is from adjusted model for age at randomization, sex, baseline NT-proBNP concentration, treatment allocation, myocardial blush grade and statin use at 4 months. Bold is *P*-value ≤ 0.0063.

Metformin decreased large LDL particles by approximately 35% without a significant effect on small LDL particles. Consequently, LDL size was also decreased. Metformin did not significantly affect plasma triglycerides, VLDL characteristics, HDL cholesterol, apoA-I and HDL subfractions. In addition, we observed that small-sized HDL particles and medium-sized VLDL obtained after 24 hours were associated with higher LVEF and a smaller infarct size.

The present findings on plasma (apo)lipoproteins and lipoprotein subfraction characteristics should be interpreted in the context of lipoprotein changes that occur in the setting of an acute coronary syndrome [10,31]. Thus, LDL cholesterol spontaneously decreases shortly after MI and rises again after several weeks [10,31]. In the current study, the LDL particle concentration was lower at 24 hours post-MI in both treatment groups, which could to at least in part be attributed by the initiation of statin therapy in the vast majority of study participants. In addition, initiation of statin treatment largely explained lower levels of total cholesterol, LDL cholesterol, non-HDL cholesterol, apoB and LDL subfractions after 4 months. These lipoprotein changes were present irrespective of metformin treatment. Plasma triglycerides may acutely decrease after an MI, followed by an increase above baseline after several days [10]. All participants received heparin before percutaneous coronary intervention mostly before arrival at the hospital as part of routine medical care [32]. It is well known that heparin stimulates lipoprotein lipase, thereby increasing lipolysis [33]. Even a low dose of heparin lowers plasma triglycerides acutely [34]. This explains our observation that plasma triglycerides were about 50% lower at presentation compared to 4 months follow-up. The VLDL particle concentration was even 5-fold lower at baseline vs, the levels obtained after 24 hours and after 4 months. On the other hand, HDL cholesterol concentration has been reported to remain fairly constant during the acute phase of MI [10,31]. Accordingly, we did not observe much change in the HDL particle concentration and in HDL subfraction levels 24 hours after manifestation of MI.

The effects of metformin when initiated during the acute phase of MI to reduce LDL cholesterol and large LDL particles has not been described in previously. In comparison, one year treatment with metformin dosed 850 mg administered twice daily decreases the LDL particle concentration, small-sized LDL particles and slightly increases LDL size along with improvement in insulin sensitivity in subjects with impaired glucose tolerance [22]. While metformin increases small HDL particles in the non-acute setting [22], we did not found a change in small HDL particles in the present study. It has also been shown that 18 months treatment with metformin at a dose of 850 mg twice daily does not significantly affect plasma total cholesterol, LDL cholesterol, HDL cholesterol and triglycerides despite improvement in insulin sensitivity in statin-using non-diabetic subjects with stable coronary heart disease [23]. These variable results emphasize the

relevance of participant selection, the circumstances of initiation of metformin treatment and possibly also of its dose and exposure time.

In the general population, an inverse relationship of HDL cholesterol with incident coronary heart disease has been consistently reported [35,36]. Moreover, the relevance of HDL subfractions for cardiovascular risk prediction has received considerable attention [6,9,12,17,28]. As yet, the importance of larger-sized compared to smaller-sized HDL particles for coronary risk has not been unequivocally established, neither in the setting of population-based cohort studies [9,12,17,28], nor in specific high risk populations [3,4,37]. Of further interest, low HDL cholesterol as determined during an acute coronary syndrome may predict recurrent cardiovascular events [38]. More recently, it was reported that HDL-associated cholesterol esterification is impaired in the acute setting of a coronary syndrome despite a lack of decrease in HDL cholesterol [39], whereas HDL anti-inflammatory function rather than HDL cholesterol may predict recurrent events [40]. In non-acute patients, impaired ability of HDL to remove cholesterol from macrophage model cells associates more closely with incident coronary heart disease than lower HDL cholesterol [41]. These findings emphasize the importance of HDL functional properties for atheroprotection [10,42]. In the context of the GIPS-III trial it is also relevant that a low HDL cholesterol concentration may represent a determinant of heart failure [43]. Experimental induction of apoA-I, HDL's most abundant apolipoprotein, improves cardiac remodeling after MI in mice [44], although HDL mimetic therapy was unsuccessful in improving cardiac outcome in humans [45]. In coronary artery disease patients, it has been cross-sectionally determined that HDL cholesterol and smaller-sized HDL may confer higher LVEF [3,46]. In the current report, we considered the metformin- and placebo-receiving participants together since metformin did not affect LVEF [27]. After adjustment for treatment allocation and other relevant covariates we demonstrated that increased concentrations of smaller-sized HDL particles prospectively predict higher LVEF. These novel results agreement with the concept that specific HDL subfractions could be pathophysiologically implicated in better cardiac performance. The mechanisms responsible for this association remain to be more precisely delineated. Among other possibilities, it could reflect the ability of certain HDL subfractions to exert anti-oxidative properties or to stimulate endothelial function [10,26,42]. In addition, higher concentrations of medium VLDL were associated with smaller infarct size. As yet the clinical implication of this association is unknown.

Several other methodological aspects of our study need to be considered. First, this randomized study was carried out in a considerable number of participants, making lack of power to demonstrate effects of metformin on VLDL and HDL subfraction characteristics as determined by NMR spectroscopy unlikely. Moreover, neither plasma triglycerides nor HDL cholesterol and apoA-I levels after 4 months of follow-up changed in response to metformin administration. Second, inherent to the design of this study to randomize

subjects shortly after arrival at the hospital, plasma lipid measurements were not carried out in the fasting state. For logistic reasons, non-fasting samples were also obtained during follow-up. However, given the placebo-controlled design of GIPS-III, it is unlikely that this approach materially affected our results. Third, only 2% of participants experienced recurrent major adverse cardiac events, and none of participants died during 4 months follow-up [27]. For this reason, associations of lipoprotein subfractions with hard clinical end-points could not be assessed. Instead, LVEF at 4 months follow-up was chosen as the pre-specified primary endpoint of the GIPS-III trial, reasoning that left ventricular dysfunction is a prevalent complication of STEMI which prospectively predicts poor cardiac outcome [26,27]. Fourth, only subjects without known diabetes participated in the GIPS-III trial [18,19]. The positive correlation between glucose at admission and MI size after 4 months, therefore, suggests that MI size may relate to stress hyperglycemia encountered during the acute phase of MI [47].

In conclusion, the present study suggests that metformin treatment initiated directly after the acute phase of MI elicits a small decrease in LDL cholesterol together with a decrease in LDL size. Furthermore, higher medium VLDL and higher small HDL particle concentrations may confer beneficial associations with increased LVEF and decreased infarct size, respectively.

REFERENCES

1. Cholesterol Treatment Trialists' (CTT) Collaboration *et al.* Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* 376, 1670–1681 (2010).
2. Perk, J. *et al.* European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Eur. Heart J.* 33, 1635–1701 (2012).
3. Kempen, H. J., van Gent, C. M., Buytenhek, R. & Buis, B. Association of cholesterol concentrations in low-density lipoprotein, high-density lipoprotein, and high-density lipoprotein subfractions, and of apolipoproteins AI and AII, with coronary stenosis and left ventricular function. *J. Lab. Clin. Med.* 109, 19–26 (1987).
4. Freedman, D. S. *et al.* Relation of lipoprotein subclasses as measured by proton nuclear magnetic resonance spectroscopy to coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* 18, 1046–1053 (1998).
5. Rosenson, R. S., Otvos, J. D. & Freedman, D. S. Relations of lipoprotein subclass levels and low-density lipoprotein size to progression of coronary artery disease in the Pravastatin Limitation of Atherosclerosis in the Coronary Arteries (PLAC-I) trial. *Am. J. Cardiol.* 90, 89–94 (2002).
6. Rosenson, R. S. *et al.* HDL measures, particle heterogeneity, proposed nomenclature, and relation to atherosclerotic cardiovascular events. *Clin. Chem.* 57, 392–410 (2011).
7. Williams, P. T., Zhao, X.-Q., Marcovina, S. M., Brown, B. G. & Krauss, R. M. Levels of cholesterol in small LDL particles predict atherosclerosis progression and incident CHD in the HDL-Atherosclerosis Treatment Study (HATS). *PloS One* 8, e56782 (2013).
8. Toth, P. P. *et al.* Cardiovascular risk in patients achieving low-density lipoprotein cholesterol and particle targets. *Atherosclerosis* 235, 585–591 (2014).
9. Joshi, P. H. *et al.* Association of high-density lipoprotein subclasses and incident coronary heart disease: The Jackson Heart and Framingham Offspring Cohort Studies. *Eur. J. Prev. Cardiol.* 2047487314543890 (2014). doi:10.1177/2047487314543890
10. Rosenson, R. S., Brewer, H. B. & Rader, D. J. Lipoproteins as biomarkers and therapeutic targets in the setting of acute coronary syndrome. *Circ. Res.* 114, 1880–1889 (2014).
11. Mackey, R. H. *et al.* High-density lipoprotein cholesterol and particle concentrations, carotid atherosclerosis, and coronary events: MESA (multi-ethnic study of atherosclerosis). *J. Am. Coll. Cardiol.* 60, 508–516 (2012).
12. Parish, S. *et al.* Lipids and lipoproteins and risk of different vascular events in the MRC/BHF Heart Protection Study. *Circulation* 125, 2469–2478 (2012).
13. Mora, S., Glynn, R. J. & Ridker, P. M. High-density lipoprotein cholesterol, size, particle number, and residual vascular risk after potent statin therapy. *Circulation* 128, 1189–1197 (2013).
14. Otvos, J. D. *et al.* Clinical implications of discordance between low-density lipoprotein cholesterol and particle number. *J. Clin. Lipidol.* 5, 105–113 (2011).
15. El Harchaoui, K. *et al.* Value of Low-Density Lipoprotein Particle Number and Size as Predictors of Coronary Artery Disease in Apparently Healthy Men and Women: The EPIC-Norfolk Prospective Population Study. *J. Am. Coll. Cardiol.* 49, 547–553 (2007).
16. Jeyarajah, E. J., Cromwell, W. C. & Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clin. Lab. Med.* 26, 847–870 (2006).

17. Mora, S. *et al.* Lipoprotein particle profiles by nuclear magnetic resonance compared with standard lipids and apolipoproteins in predicting incident cardiovascular disease in women. *Circulation* 119, 931–939 (2009).
18. Mora, S. *et al.* LDL particle subclasses, LDL particle size, and carotid atherosclerosis in the Multi-Ethnic Study of Atherosclerosis (MESA). *Atherosclerosis* 192, 211–217 (2007).
19. Brunzell, J. D. *et al.* Lipoprotein Management in Patients With Cardiometabolic Risk Consensus statement from the American Diabetes Association and the American College of Cardiology Foundation. *Diabetes Care* 31, 811–822 (2008).
20. Arsenault, B. J. *et al.* Lipid assessment, metabolic syndrome and coronary heart disease risk. *Eur. J. Clin. Invest.* 40, 1081–1093 (2010).
21. Mora, S., Buring, J. E. & Ridker, P. M. Discordance of low-density lipoprotein (LDL) cholesterol with alternative LDL-related measures and future coronary events. *Circulation* 129, 553–561 (2014).
22. Goldberg, R. *et al.* Lifestyle and Metformin Treatment Favorably Influence Lipoprotein Subfraction Distribution in the Diabetes Prevention Program. *J. Clin. Endocrinol. Metab.* 98, 3989–3998 (2013).
23. Preiss, D. *et al.* Metformin for non-diabetic patients with coronary heart disease (the CAMERA study): a randomised controlled trial. *Lancet Diabetes Endocrinol.* 2, 116–124 (2014).
24. Trifunovic, D. *et al.* Acute insulin resistance in ST-segment elevation myocardial infarction in non-diabetic patients is associated with incomplete myocardial reperfusion and impaired coronary microcirculatory function. *Cardiovasc. Diabetol.* 13, 73 (2014).
25. Sanjuan, R. *et al.* Insulin resistance and short-term mortality in patients with acute myocardial infarction. *Int. J. Cardiol.* 172, e269-270 (2014).
26. Lexis, C. P. H. *et al.* Metformin in non-diabetic patients presenting with ST elevation myocardial infarction: rationale and design of the glycometabolic intervention as adjunct to primary percutaneous intervention in ST elevation myocardial infarction (GIPS)-III trial. *Cardiovasc. Drugs Ther. Spons. Int. Soc. Cardiovasc. Pharmacother.* 26, 417–426 (2012).
27. Lexis, C. P. H. *et al.* Effect of metformin on left ventricular function after acute myocardial infarction in patients without diabetes: the GIPS-III randomized clinical trial. *JAMA* 311, 1526–1535 (2014).
28. Oksala, N. *et al.* Complementary prediction of cardiovascular events by estimated apo- and lipoprotein concentrations in the working age population. The Health 2000 Study. *Ann. Med.* 45, 141–148 (2012).
29. Wong, D. T. L. *et al.* The role of cardiac magnetic resonance imaging following acute myocardial infarction. *Eur. Radiol.* 22, 1757–1768 (2012).
30. van 't Hof, A. W. *et al.* Angiographic assessment of myocardial reperfusion in patients treated with primary angioplasty for acute myocardial infarction: myocardial blush grade. Zwolle Myocardial Infarction Study Group. *Circulation* 97, 2302–2306 (1998).
31. Avogaro, P. *et al.* Variations in apolipoproteins B and A, during the course of myocardial infarction. *Eur. J. Clin. Invest.* 8, 121–129 (1978).
32. Bouma, M., Rutten, F. H., Wiersma, T. & Burgers, J. S. [Revised Dutch College of General Practitioners' practice guideline 'Acute coronary syndrome']. *Ned. Tijdschr. Geneeskd.* 157, A6006 (2013).
33. Riemens, S. C., van Tol, A., Scheek, L. M. & Dullaart, R. P. Plasma cholesteryl ester transfer and hepatic lipase activity are related to high-density lipoprotein cholesterol in association with insulin resistance in type 2 diabetic and non-diabetic subjects. *Scand. J. Clin. Lab. Invest.* 61, 1–9 (2001).
34. Brunner, M. P. *et al.* Effect of heparin administration on metabolomic profiles in samples obtained during cardiac catheterization. *Circ. Cardiovasc. Genet.* 4, 695–700 (2011).

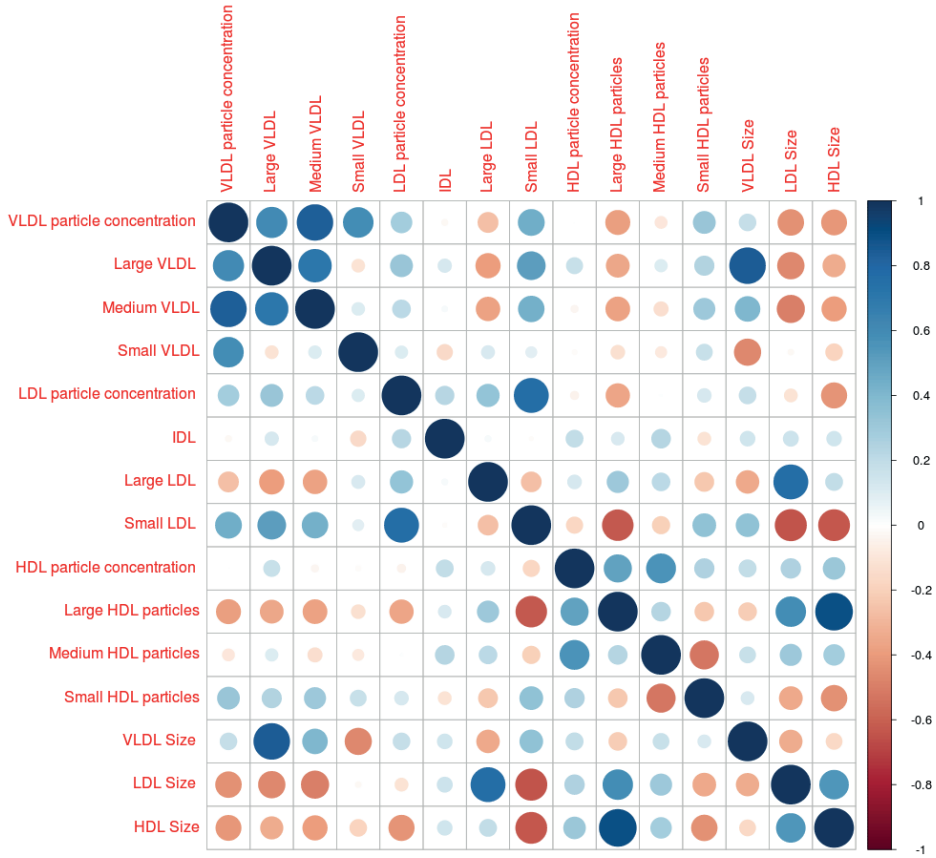
35. Prospective Studies Collaboration *et al.* Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet Lond. Engl.* 370, 1829–1839 (2007).
36. The Emerging Risk Factors Collaboration*. MAjor lipids, apolipoproteins, and risk of vascular disease. *JAMA* 302, 1993–2000 (2009).
37. Xu, R.-X. *et al.* High-density lipoprotein subfractions in relation with the severity of coronary artery disease: A Gensini score assessment. *J. Clin. Lipidol.* 9, 26–34 (2015).
38. Olsson, A. G. *et al.* High-density lipoprotein, but not low-density lipoprotein cholesterol levels influence short-term prognosis after acute coronary syndrome: results from the MIRACL trial. *Eur. Heart J.* 26, 890–896 (2005).
39. Dullaart, R. P. F. *et al.* Alterations in plasma lecithin:cholesterol acyltransferase and myeloperoxidase in acute myocardial infarction: implications for cardiac outcome. *Atherosclerosis* 234, 185–192 (2014).
40. Dullaart, R. P. F., Annema, W., Tio, R. A. & Tietge, U. J. F. The HDL anti-inflammatory function is impaired in myocardial infarction and may predict new cardiac events independent of HDL cholesterol. *Clin. Chim. Acta Int. J. Clin. Chem.* 433, 34–38 (2014).
41. Khera, A. V. *et al.* Cholesterol efflux capacity, high-density lipoprotein function, and atherosclerosis. *N. Engl. J. Med.* 364, 127–135 (2011).
42. Triolo, M., Annema, W., Dullaart, R. P. F. & Tietge, U. J. F. Assessing the functional properties of high-density lipoproteins: an emerging concept in cardiovascular research. *Biomark. Med.* 7, 457–472 (2013).
43. Velagaleti, R. S. *et al.* Relations of lipid concentrations to heart failure incidence: the Framingham Heart Study. *Circulation* 120, 2345–2351 (2009).
44. Gordts, S. C. *et al.* Beneficial effects of selective HDL-raising gene transfer on survival, cardiac remodelling and cardiac function after myocardial infarction in mice. *Gene Ther.* 20, 1053–1061 (2013).
45. Tardif, J.-C. *et al.* Effects of the high-density lipoprotein mimetic agent CER-001 on coronary atherosclerosis in patients with acute coronary syndromes: a randomized trial. *Eur. Heart J.* 35, 3277–3286 (2014).
46. Yang, N. *et al.* Variability in lipid profile among patients presented with acute myocardial infarction, unstable angina and stable angina pectoris. *Eur. Rev. Med. Pharmacol. Sci.* 18, 3761–3766 (2014).
47. Fujino, M. *et al.* Impact of acute and chronic hyperglycemia on in-hospital outcomes of patients with acute myocardial infarction. *Am. J. Cardiol.* 114, 1789–1793 (2014).

Supplementary Appendix

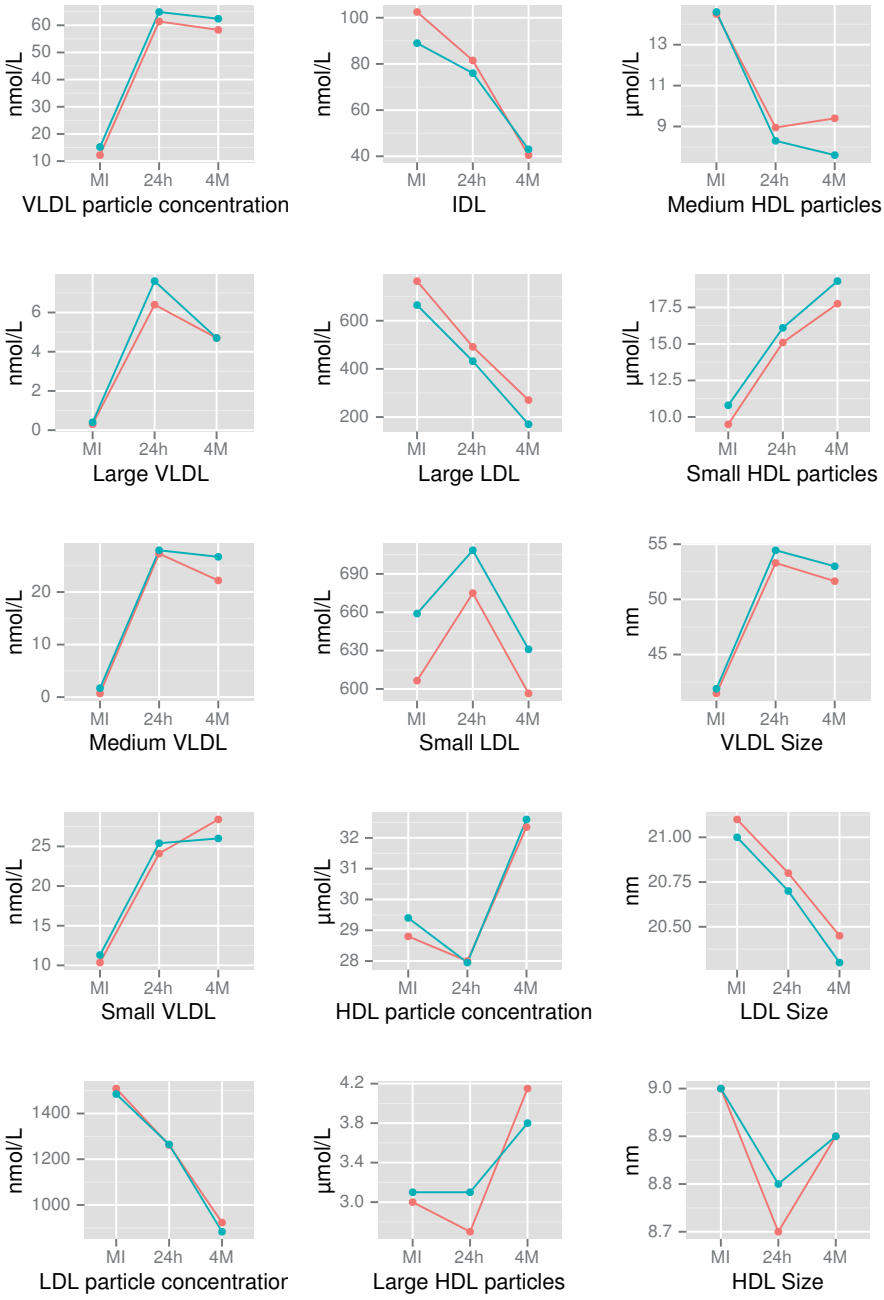
Supplementary Figure 1 | Correlation matrix using Pearson correlation coefficients from the metabolite concentrations of the lipoprotein subfractions 4 months post-MI (n = 317). 117

Supplementary Figure 2 | Panel showing changes in median of metabolite levels over the three time points Myocardial Infarction (MI), 24 hours post-MI (24h) and 4 months post-MI (4M) on the X- axis. 118

Supplementary Table 1 | Overview of Principal Components (PC) analysis of the lipoprotein subfractions showing the importance of the components. 119



Supplementary Figure 1. Correlation matrix using Pearson correlation coefficients from the metabolite concentrations of the lipoprotein subfractions 4 months post-MI (n = 317)
Data is shown for the discovery analysis of N:134,251 individuals excluding genetic variants with minor allele frequency of <0.001, and information measure <0.3 leaving 19,941,912 genetic variants.



Supplementary Figure 2. Panel showing changes in median of metabolite levels over the three time points Myocardial Infarction (MI), 24 hours post-MI (24h) and 4 months post-MI (4M) on the X- axis. Red lines represent placebo and blue lines metformin treatment. The Y-axis represents the metabolite concentration with units given in either nmol/L, µmol/L, or nm.

Supplementary Table 1. Overview of Principal Components (PC) analysis of the lipoprotein subfractions showing the importance of the components

Importance of components:	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**	**PC10**
Standard deviation	24.196	16.774	13.556	105.116	10.122	0.83646	0.75975	0.72186	0.44544	0.34392
Proportion of Variance	0.3903	0.1876	0.1225	0.07366	0.0683	0.04664	0.03848	0.03474	0.01323	0.00789
Cumulative Proportion	0.3903	0.5779	0.7004	0.77406	0.8424	0.88900	0.92748	0.96222	0.97545	0.98333
Importance of components:	**PC11**	**PC12**	**PC13**	**PC14**	**PC15**					
Standard deviation	0.29727	0.28863	0.18338	0.17469	0.11910					
Proportion of Variance	0.00589	0.00555	0.00224	0.00203	0.00095					
Cumulative Proportion	0.98922	0.99478	0.99702	0.99905	100.000					

Chapter 6

Effect of Metformin on Metabolites and Relation with Myocardial Infarct Size and Left Ventricular Ejection Fraction After Myocardial Infarction

Ruben N. Eppinga*, Daniel Kofink*, Robin P.F. Dullaart, Geertje W. Dalmeijer, Erik Lipsic, Dirk J. van Veldhuisen, Iwan C.C. van der Horst, Folkert W. Asselbergs†, Pim van der Harst†

* both authors contributed equally, † both authors contributed equally

Adapted from Circ Cardiovasc Genet. 2017 Feb;10(1). pii: e001564

ABSTRACT

Background: Left ventricular ejection fraction (LVEF) and infarct size (ISZ) are key predictors of long-term survival after myocardial infarction (MI). However, little is known about the biochemical pathways driving left ventricular dysfunction after MI. To identify novel biomarkers predicting post-MI LVEF and ISZ, we performed metabolic profiling in the GIPS-III randomized clinical trial. We also investigated the metabolic footprint of metformin, a drug associated with improved post-MI left ventricular function in experimental studies.

Methods and results: Participants were ST-elevated MI (STEMI) patients who were randomly assigned to receive metformin or placebo for 4 months. Blood samples were obtained on admission, 24 h and 4 months post-MI. 233 metabolite measures were quantified using nuclear magnetic resonance (NMR) spectrometry. LVEF and ISZ were assessed 4 months post-MI. 24 h post-MI measurements of HDL triglycerides (HDL-TG) predicted LVEF (β=1.90 [95% CI: 0.82, 2.98]; $p=6.4x10^{-4}$) and ISZ (β=-0.41; 95% CI: -0.60, -0.21]; $p=3.2x10^{-5}$). Additionally, 24 h post-MI measurements of medium HDL-TG (β=-0.40 [95% CI: -0.60, -0.20]; $p=6.4x2x10^{-5}$), small HDL-TG (β=-0.34 [95% CI: -0.53, -0.14]; $p=7.3x10^{-4}$) and the triglyceride content of very large HDL (β=-0.38 [95% CI: -0.58, -0.18]; $p=2.7x10^{-4}$) were associated with ISZ. After the 4-month treatment, the phospholipid content of very large HDL was lower in metformin vs. placebo treated patients (28.89% vs. 38.79%; $p=7.5x10^{-5}$); alanine levels were higher in the metformin group (0.46 mmol/L vs. 0.44 mmol/L; $p=2.4x10^{-4}$).

Conclusions: HDL triglyceride concentrations predict post-MI LVEF and ISZ. Metformin increases alanine levels and reduces the phospholipid content in very large HDL particles.

INTRODUCTION

Myocardial infarction (MI) is one of the leading causes of global morbidity and mortality. While the survival after MI has improved due to ameliorated treatment strategies, including primary percutaneous interventions, the long-term outcome of MI in general remains poor with a 1-year risk for recurrent cardiovascular (CV) disease of over 10%.[1] Left ventricular ejection fraction (LVEF) and infarct size (ISZ) are key predictors of long-term prognosis after MI.[2,3]. However, treatment options for left ventricular dysfunction are limited and the biochemical mechanisms driving functional decline of the myocardium after MI are largely unknown.

Metformin, which is commonly used in the treatment of diabetes and more recently in insulin resistant conditions, has been found to preserve LVEF and to reduce ISZ in non-diabetic animal models of MI.[4] The GIPS-III clinical trial was designed to study the effects of metformin therapy on LVEF in non-diabetic ST segment Elevation MI (STEMI) patients undergoing PCI. However, in contrast to preclinical findings, metformin did not improve LVEF compared with placebo 4 months post-MI.[5]

This result may be explained by interindividual differences in metformin response, raising the possibility that metformin is effective in a subgroup of CV patients. Metabolic profiling has emerged as a powerful tool to explore drug effects and factors influencing drug response.[6-8] Metabolomics is a relatively novel field in 'omics' sciences, which uses high-throughput technologies, such as nuclear magnetic resonance (NMR) spectroscopy, to concurrently quantify a large number of small molecules in different tissues. While recent studies reported changes in lipid and amino acid concentrations after metformin treatment[9-11], no study has yet used large-scale metabolic platforms to investigate the effects of metformin on a wide range of metabolite measures at a time. Furthermore, metabolic profiling has been performed to improve diagnosis and prediction of CV events.[12,13] A recent study identified metabolic profiles which discriminate heart failure patients from healthy controls.[14] Metabolic profiling may thus help identify novel biomarkers of left ventricular function and ISZ to improve risk stratification in MI patients.

Metabolite concentrations can vary greatly over time and are highly sensitive to environmental influences. Lipid profiles have been shown to change shortly after MI and only gradually return to baseline after several weeks.[15] The predictive value of a biomarker may thus vary over time. We therefore studied metabolic markers of LVEF, ISZ and metformin response in the GIPS-III cohort at three different time points: baseline (on admission), 24 h post-MI and 4 months post-MI.

The objective of this ancillary study of the GIPS-III trial was to evaluate the effect of metformin on metabolic profiles in non-diabetic STEMI patients and to identify prognostic markers, which predict LVEF and ISZ 4 months post-MI. Furthermore we tested whether metformin improved LVEF and ISZ in subgroups of patients, as identified by metabolic profiling.

METHODS

Study population

The GIPS-III study is a randomized trial that included 380 non-diabetic patients undergoing primary PCI for ST segment elevation myocardial infarction. Participants received a 4-month regimen with either metformin 500mg 2dd1 or matching placebo 2dd1. The design of the study has been previously described in more detail.[4,5] All patients provided written informed consent. The study complied with the Declaration of Helsinki and was approved by the ethics committee of the University Medical Center Groningen (the Netherlands) and national authorities (NCT01217307). The primary outcome measure was LVEF, the secondary outcome measure was ISZ.

Laboratory measurements

Non-fasting blood samples were obtained on admission (N=339), 24 h post-MI (N=329) and 4 months post-MI (N=316). Serum and EDTA-anticoagulated plasma samples were stored at – 80 °C until analyzed. Metabolic profiling was performed using a high-throughput ^{1}H NMR metabolomics platform.[16] We obtained a total of 233 serum metabolite concentrations and ratios, including 168 lipoprotein subclass measures, 45 lipid related measures, 5 glycolysis related metabolites, 9 amino acids, 3 ketone bodies, 2 fluid balance related metabolites and 1 inflammatory marker. An overview of all metabolite measures is given in Supplemental Table 1.

Cardiac Magnetic Resonance imaging (CMR)

LVEF and ISZ were measured by cardiac magnetic resonance imaging (MRI) 4 months after MI as previously described in detail.[4,5]. Independent cardiologists analyzed all MRI data and assessed LVEF and ISZ, blinded for treatment assignment.

Statistical analysis

Missing metabolite measures were imputed using random forest imputation as implemented in the R package missForest.[17] Since most metabolite measures showed skewed distributions, they were normalized using rank-based inverse normal transformation within each time point separately. Spearman's correlation coefficients were calculated from the metabolite concentrations for each time point (baseline, 24 h post-MI and 4 months post-MI) and plotted using the corrplot function of the corrplot package of R. The correlation plots are presented in Supplemental Figure 1.

Since many metabolites were highly correlated, principal component analysis (PCA) was applied to estimate the number of independent tests for multiple testing correction, using the prcomp-function in R. To additionally account for multiple testing at different time points, principal components (PCs) were calculated across all three time points.

The first 68 PCs explained over 95% of the variation in the metabolite data, yielding an adjusted significance level of $p<0.05/68 = 0.00074$.

Unpaired t-tests were performed to assess the effect of metformin treatment on metabolite measures. To identify biomarkers predictive of LVEF and ISZ, we analyzed all metabolite measures at each time point separately, using linear regression adjusted for known predictors of ventricular function and medication use: age, sex, baseline N-terminal prohormone of brain natriuretic peptide (NTproBNP) levels, baseline creatine kinase (CK)-MB levels, myocardial blush grade, metformin treatment and statin treatment (4 months post-MI). To meet the assumption of normality of residuals, we tested different transformations. Since square-root transformation provided the best results, ISZ was square-root transformed. In addition we performed stratified analyses for LVEF and ISZ. According to current guidelines[18], LVEF 52%-72% was categorized as normal ventricular function; LVEF 41-51% was defined as mildly abnormal and LVEF<41% as abnormal for men. Categories were LVEF 54%-74% for normal ventricular function, LVEF 41-53% as mildly abnormal and LVEF<41% for abnormal for women. ISZ was stratified by tertiles to obtain the same number of strata as with LVEF. Associations of metabolite measures with LVEF categories and ISZ tertiles were assessed using multinomial logistic regression, which provides pairwise comparisons between each level of the outcome variable and a reference level. Finally we added the interaction term of metformin treatment and metabolite measure to the linear regression models to identify subgroups of patients in whom metformin was effective. R (version 3.02 or higher, http://www.r-project.org/) was used for all statistical analyses.

RESULTS

Patient characteristics and metabolite measures

A total of 380 patients received either metformin placebo treatment. Of these, 109 did not undergo MRI 4 months post-MI or did not provide utilizable scans due to insufficient quality. Details on metformin/placebo treatment, clinical parameters and conventional lipid and (apo)lipoprotein measures have been published elsewhere.[5,11] Briefly, metformin treatment resulted in a modest decrease in low-density lipoprotein cholesterol (LDL-C) without significant effects on total cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides, apolipoprotein B (apoB) and apolipoprotein A-I (apoA1) when the values after 4 months and after 24 h were compared (data not shown).[14] Metabolic profiles were quantified in a total of 376 patients. Baseline, 24 h post-MI and 4-month post-MI measurements were available from 339, 326 and 316 patients, respectively. Premature dropout was neither related to metformin treatment, nor to mortality as none of the participants died before MRI.[5] A summary of all metabolite measures can be found

in Supplemental Table 1. The correlation matrices revealed substantial correlation within lipoprotein subclasses, between amino acids and between fatty acids (Supplemental Figure 1).

Association of metabolites with LVEF and ISZ

Results for all metabolite measures tested are shown in Supplemental Tables 2-7. None of the metabolite measures was significantly associated with LVEF 4 months post-MI. No baseline metabolite measure predicted LVEF. Patients with higher HDL-TG levels 24 h post-MI showed significantly better LVEF (β=1.90 [95% CI: 0.82, 2.97]; $p=6.4x10^{-4}$) after adjustment for metformin treatment, age, sex, baseline NTproBNP levels, baseline CK-MB levels, myocardial blush grade and statin use (Table 1). When LVEF was entered as categorical variable (normal, mildly abnormal, abnormal ventricular function), 24 h post-MI measurements of HDL-TG (OR=0.36 [95% CI: 0.21, 0.61]; $p=1.8x10^{-4}$), medium (M-) HDL-TG (OR=0.37 [95% CI: 0.22, 0.63]; $p=2.3x10^{-4}$) and small (S-) HDL-TG (OR=0.35 [95% CI: 0.20, 0.61]; $p=2.1x10^{-4}$) significantly predicted normal vs. abnormal LVEF 4 months

Table 1. Association of selected metabolite measures with LVEF and ISZ 24 h post-MI

	Unadjusted model		Adjusted model	
Metabolite	**β [95% CI]**	**p**	**β [95% CI]**	**p**
LVEF (N=245)				
HDL-TG	1.84 [0.78, 2.89]	$7.4x10^{-4}$	1.90 [0.82, 2.97]	$6.4x10^{-4*}$
M-HDL-TG	1.70 [0.65, 2.75]	0.002	1.65 [0.55, 2.74]	0.004
XL-HDL-TG%	1.67 [0.56, 2.77]	0.003	1.82 [0.68, 2.96]	0.002
S-HDL-TG	1.51 [0.45, 2.57]	0.006	1.68 [0.58, 2.78]	0.003
Albumin	1.09 [0.04, 2.16]	0.044	1.25 [0.10, 2.40]	0.034
Phenylalanine	-0.90 [-2.01, 0.21]	0.113	-0.55 [-1.68, 0.58]	0.344
ISZ (N=231)				
HDL-TG	-0.42 [-0.60, -0.24]	$1.2x10^{-5*}$	-0.41 [-0.60, -0.22]	$3.2x10^{-5*}$
M-HDL-TG	-0.42 [-0.60, -0.23]	$1.4x10^{-5*}$	-0.40 [-0.60, -0.21]	$6.4x10^{-5*}$
XL-HDL-TG%	-0.37 [-0.56, -0.18]	$1.9x10^{-4*}$	-0.38 [-0.58, -0.18]	$2.7x10^{-4*}$
S-HDL-TG	-0.33 [-0.52, -0.14]	$6.8x10^{-4*}$	-0.34 [-0.54, -0.15]	$7.3x10^{-4*}$
Albumin	-0.33 [-0.52, -0.15]	$5.2x10^{-4*}$	-0.33 [-0.54, -0.13]	0.002
Phenylalanine	0.38 [0.18, 0.58]	$1.9x10^{-4*}$	0.34 [0.14, 0.55]	0.001

Results are shown for the unadjusted model and the adjusted model including age, sex, treatment, statin use, CKMB, NTproBNP and MBG as covariates. *Effects significant after correction for multiple testing ($p<7.4x10^{-4}$). CI: confidence interval; LVEF: left ventricular ejection fraction; ISZ: infarct size; HDL-TG: triglycerides in HDL particles; M-HDL-TG: triglycerides in medium HDL particles; XL-HDL-TG%: triglycerides to total lipids ratio in very large HDL particles; S-HDL-TG: triglycerides in small HDL particles.

post-MI (Figure 1A, Table 2). Notably, all HDL-TG related measures showed a positive association with LVEF, suggesting a beneficial effect of increased triglyceride content in HDL. We found no association of 24-h post-MI measurements with mildly abnormal LVEF relative to normal LVEF. In addition, 24 h post-MI measurements of triglycerides (OR=0.39 [95% CI: 0.23, 0.66]; $p=5.2x10^{-4}$) and the cholesterol (OR=2.52 [95% CI: 1.48, 4.30]; $p=6.6x10^{-4}$) in very small (XS-) very low-density lipoprotein (VLDL) particles was associated with abnormal LVEF compared to normal left ventricular function. Finally, baseline measurements of the TG to total lipids ratio in large (L) LDL predicted abnormal LVEF (OR=0.37 [95% CI: 0.21, 0.65]; $p=6.2x10^{-4}$). Addition of a treatment x metabolite interaction term did not reveal any patient subgroup in whom metformin improved LVEF (Table 3).

Table 2. Associations of selected NMR measures with LVEF and ISZ categories

Metabolite	**OR [95% CI]**	**p**	**OR [95% CI]**	**p**
LVEF	normal vs. mildly abnormal		normal vs. abnormal	
Baseline				
L-LDL-TG%	1.05 [0.76, 1.45]	0.774	0.37 [0.21, 0.65]	$6.2x10^{-4*}$
24 h post-MI				
HDL-TG	0.70 [0.51, 0.96]	0.028	0.36 [0.21, 0.61]	$1.8x10^{-4*}$
M-HDL-TG	0.72 [0.52, 1.00]	0.050	0.37 [0.22, 0.63]	$2.3x10^{-4*}$
S-HDL-TG	0.74 [0.53, 1.02]	0.062	0.35 [0.20, 0.61]	$2.1x10^{-4*}$
XS-VLDL-TG%	0.83 [0.61, 1.14]	0.247	0.39 [0.23, 0.66]	$5.2x10^{-4*}$
XS-VLDL-C%	1.10 [0.82, 1.49]	0.523	2.52 [1.48, 4.30]	$6.6x10^{-4*}$
ISZ	1st tertile vs. 2nd tertile		1st tertile vs. 3rd tertile	
24 h post-MI				
HDL-TG	0.78 [0.55, 1.11]	0.169	0.48 [0.33, 0.69]	$9.2x10^{-5*}$
M-HDL-TG	0.88 [0.62, 1.25]	0.472	0.46 [0.31, 0.67]	$6.2x10^{-5*}$
S-HDL-TG	0.92 [0.64, 1.30]	0.620	0.51 [0.35, 0.74]	$3.9x10^{-4*}$
XL-HDL-TG%	0.75 [0.52, 1.07]	0.116	0.49 [0.33, 0.72]	$3.2x10^{-4*}$

Associations of metabolite measures 24 h post-MI with LVEF categories (normal, mildly abnormal, abnormal) and ISZ (tertiles) categories, adjusted for age, sex, treatment, statin use, CKMB, NTproBNP and MBG. Results for pairwise comparisons are given. Metabolites with at one significant pairwise between-group comparison are shown. Effects significant after correction for multiple testing ($p<7.4x10^{-4}$) are highlighted in bold. CI: confidence interval; LVEF: left ventricular ejection fraction; ISZ: infarct size; L-LDL-TG%: triglyceride to total lipids ratio in large LDL particles; HDL-TG: triglycerides in HDL particles; M-HDL-TG: triglycerides in medium HDL particles; S-HDL-TG: triglycerides in small HDL particles; XL-HDL-TG%: triglycerides to total lipids ratio in very large HDL particles; XS-VLDL-TG%: triglycerides to total lipids ratio in very small VLDL particles; XS-VLDL-C%: cholesterol to total lipids ratio in very small VLDL particles.

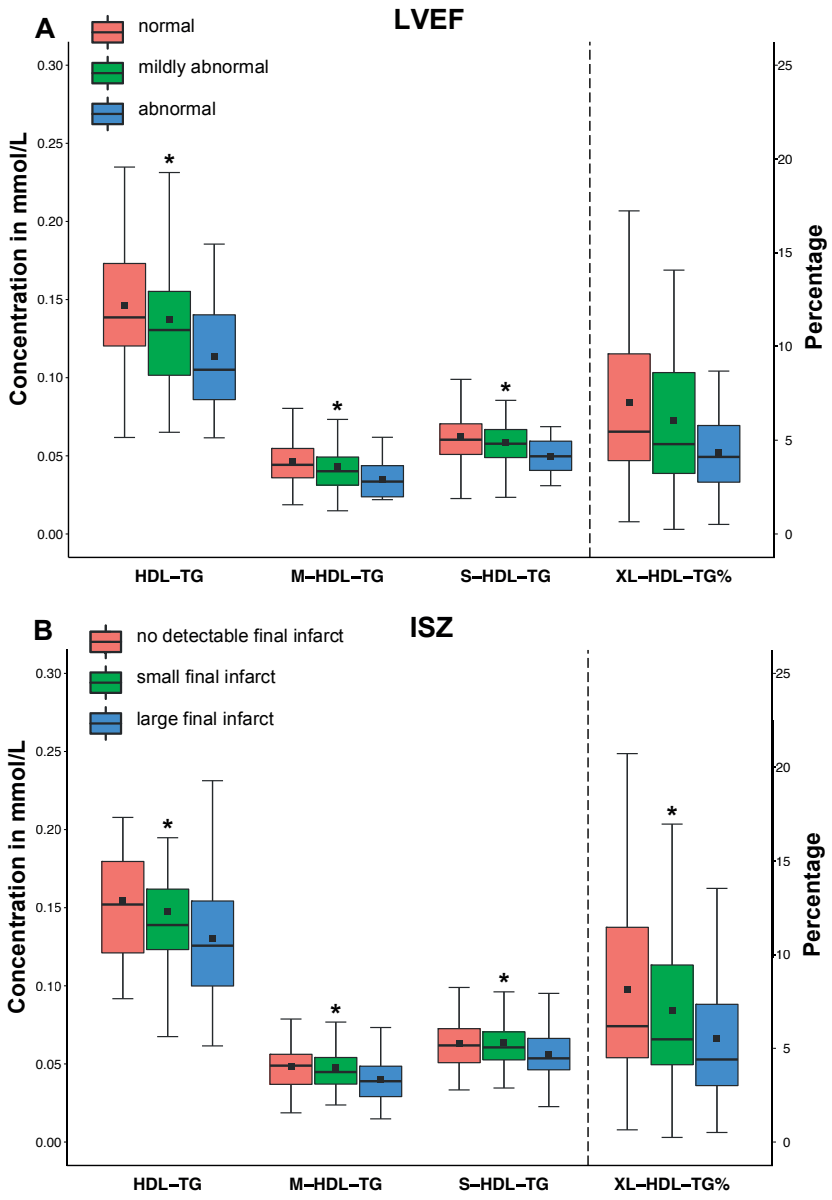


Figure 1. Categorical analysis for LVEF and ISZ

Box plots comparing selected NMR measures (24 h post-MI) between distinct LVEF (**A**) and ISZ (**B**) categories. For all plots, the lower and the upper margins represent the first and third quartile, respectively. Vertical lines indicate median values; squares indicate mean values. The whiskers represent the lowest and the highest value within 1.5 interquartile ranges. Outliers are not shown. Differences between categories were assessed using multinomial logistic regression adjusted for treatment, age, sex, NTproBNP levels, CK-MB levels, myocardial blush grade, statin use. *Effects significant after correction for multiple testing ($p<7.4x10-4$). Abbreviations are as in Table 2.

Table 3. Association of treatment x metabolite interaction for selected metabolites

	Unadjusted model		Adjusted model	
Metabolite	**β [95% CI]**	**p**	**β [95% CI]**	**p**
LVEF (N=245)				
Treatment X				
HDL-TG	1.15 [-0.96, 3.26]	0.285	0.94 [-1.18, 3.06]	0.387
M-HDL-TG	1.09 [-1.01, 3.19]	0.311	0.82 [-1.30, 2.94]	0.449
XL-HDL-TG%	0.71 [-1.50, 2.92]	0.529	0.60 [-1.61, 2.82]	0.593
S-HDL-TG	0.73 [-1.42, 2.87]	0.506	0.68 [-1.46, 2.82]	0.534
Albumin	0.36 [-1.77, 2.49]	0.740	0.38 [-1.74, 2.51]	0.725
Phenylalanine	-1.02 [-3.25, 1.22]	0.373	-1.24 [-3.51, 1.03]	0.286
ISZ (N=231)				
Treatment X				
HDL-TG	-0.07 [-0.44, -0.30]	0.701	-0.04 [-0.41, 0.34]	0.856
M-HDL-TG	-0.19 [-0.56, 0.18]	0.316	-0.14 [-0.52, 0.24]	0.465
XL-HDL-TG%	0.04 [-0.35, 0.42]	0.849	0.04 [-0.35, 0.44]	0.827
S-HDL-TG	-0.04 [-0.42, 0.35]	0.857	-0.03 [-0.41, 0.36]	0.891
Albumin	-0.06 [-0.43, 0.32]	0.762	-0.05 [-0.43, 0.33]	0.804
Phenylalanine	0.01 [-0.39, 0.40]	0.977	0.04 [-0.36, 0.45]	0.834

Results are shown for the unadjusted model (main effects and interaction term) and adjusted model including age, sex, statin use, CKMB, NTproBNP and MBG as covariates. CI: confidence interval; LVEF: left ventricular ejection fraction; ISZ: infarct size; HDL-TG: triglycerides in HDL particles; M-HDL-TG: triglycerides in medium HDL particles; XL-HDL-TG%: triglycerides to total lipids ratio in very large HDL particles; S-HDL-TG: triglycerides in small HDL particles.

We did not find any association between metabolite measures and ISZ at baseline and 4 months post-MI. In the adjusted model, HDL-TG (β=-0.41 [95% CI: -0.60, -0.22]; $p=3.2x10^{-5}$), M-HDL-TG (β=-0.40 [95% CI: -0.60, -0.21]; $p=6.42x10^{-5}$), XL-HDL-TG% (β=-0.38 [95% CI: -0.58, -0.18]; $p=2.7x10^{-4}$) and S-HDL-TG (β=-0.34 [95% CI: -0.54, -0.15]; $p=7.3x10^{-4}$) were significantly associated with ISZ 24 h post-MI (Table 1). In addition, phenylalanine (β=0.38 [95% CI: 0.18, 0.58]; $p=1.9x10^{-4}$) and albumin (β=-0.33 [95% CI: -0.52, -0.15]; $p=5.2x10^{-4}$) reached significance in the unadjusted model, but not in the adjusted model. Similarly, 24 h post-MI measurements of HDL-TG (OR=0.48 [95% CI: 0.33, 0.69]; $p=9.2x10^{-5}$), M-HDL-TG (OR=0.46 [95% CI: 0.31, 0.67]; $p=6.2x10^{-5}$), S-HDL-TG (OR=0.51 [95% CI: 0.35, 0.74]; $p=3.9x10^{-4}$) and XL-HDL-TG% (OR=0.49 [95% CI: 0.33, 0.72]; $p=3.2x10^{-4}$) predicted ISZ, when the first tertile was compared to the third tertile (Figure 1B, Table 2). Again, our findings suggest a beneficial effect of higher HDL-TG levels. We found no significant treatment x metabolite interactions (Table 3). As shown in Supplemental Table 8, HDL-TG, M-HDL-TG, S-HDL-TG and XL-HDL-TG% increased

Table 4. Treatment effects of metformin

	Placebo	Metformin	
Metabolite	**Median (IQR)**	**Median (IQR)**	**p**
24 h post-MI	N=170	N=159	
Alanine in mmol/l	0.46 (0.09)	0.49 (0.09)	$9.0x10^{-4}$
Pyridine in mmol/l	0.14 (0.05)	0.16 (0.07)	0.001
XL-HDL-PL%	36.11 (17.52)	33.98 (16.65)	0.908
4 months post-MI	N=159	N=157	
Alanine in mmol/l	0.44 (0.08)	0.46 (0.09)	$2.4x10^{-4*}$
Pyridine in mmol/l	0.1 (0.04)	0.11 (0.04)	0.006
XL-HDL-PL%	38.79 (19.5)	29.89 (23.9)	$7.5x10^{-5*}$

Effects of treatment on selected metabolite measures 24 h post-Myocardial Infarction (MI) and 4 months post-MI. *Significant effects ($p<7.4x10^{-4}$). IQR: inter-quartile range; XL-HDL-PL%: phospholipids to total lipids ratio in very large HDL particles.

between baseline and 24 h post-MI and remained relatively stable between 24 h and 4 months post-MI, except for XL-HDL-TG% which showed a moderate gain. Similar to HDL-TG, serum triglyceride levels increased between baseline and 24 h post-MI, but were decreased 4 months after MI.

Effects of metformin of metabolic profiles

Results for all metabolite measures are shown in Supplemental Table 9. To assess baseline differences in metabolic profiles, we compared metabolite measures between the treatment group and controls at baseline. We did not find any difference between the two groups at baseline. Table 4 summarizes metabolic measurements for 24 h post-MI and 4 months post-MI. 24 h post-MI, after the first doses of the treatment had been administered, both alanine (median: 0.49 mmol/L vs. 0.46 mmol/L; $p=9.0x10^{-4}$) and pyruvate (median: 0.16 mmol/L vs. 0.14 mmol/L; p=0.001) displayed trends towards increased concentrations in the metformin group. After the 4-month treatment period, alanine levels were significantly elevated in metformin-treated patients (median: 0.46 mmol/L vs. 0.44 mmol/L; $p=2.4x10^{-4}$). In addition, the phospholipids to total lipids ratio in very large high density lipoprotein (XL-HDL) particles (XL-HDL-PL%) was significantly reduced in the metformin group (median: 28.89% vs. 38.79%; $p=7.5x10^{-5}$).

DISCUSSION

We used ^{1}H NMR spectrometry-based metabolite measures to evaluate the effects of metformin on metabolic profiles of non-diabetic MI patients and to study prognostic metabolites predicting LVEF and ISZ 4 months post-MI. Moreover we investigated

whether metabolic profiling could be used to identify subgroups of patients in whom metformin was effective. After the 4-month treatment period, we found higher alanine levels and lower XL-HDL-PL% in metformin-treated patients as compared to controls. Remarkably, higher triglyceride levels in HDL and several HDL subfractions measured 24 h post-MI were associated with favorable outcome as inferred from higher LVEF and smaller ISZ 4 months post-MI. Moreover, categorical analysis of LVEF revealed that besides HDL-TG, the composition of XS-VLDL (24 h post-MI) and L-LDL (baseline) was associated with abnormal left ventricular function 4 months post-MI. We could not identify metabolic profiles associated with treatment benefits from metformin.

Similar to our results, the CAMERA study, a clinical trial investigating the effects of metformin on different amino acids, found substantially increased alanine levels 18 months after treatment onset.[13] Alanine plays a crucial role in the alanine-glucose cycle, in which alanine released by muscle tissue is transported to the liver before it is converted into pyruvate for gluconeogenesis. Findings from animal studies suggest that metformin reduces gluconeogenesis by inhibiting hepatic alanine uptake[19] and by hampering fat-induced changes in the glycolysis metabolic pathway[20]. As a result of reduced uptake into the liver, blood alanine levels may rise in metformin-treated patients. Interestingly, we observed a trend towards increased alanine levels in the metformin group 24 h post-MI,,suggesting rapid effects of metformin on gluconeogenesis.

Numerous randomized controlled trials have studied the effects of metformin treatment on lipid levels in patients with type 2 diabetes. A recent study in diabetic patients found that metformin lowered total cholesterol and LDL-C.[21] Another study in patients at risk for diabetes reported changes in lipoprotein subclasses after one year of metformin treatment, with reduced particle concentrations of small LDL and elevated large LDL, small HDL and large HDL.[9] In our recent report, we observed modest decreases in LDL cholesterol, no change in apolipoprotein B, and as a result a small decrease in LDL particle size.[11] In the present study which used a different NMR-based method, only the phospholipid content of large HDL particles was decreased in response to metformin.

We also tested whether lipoprotein characteristics and metabolite measures at baseline, 24 h post-MI and 4 months post-MI were associated with 4 months post-MI LVEF and ISZ. We found that increased HDL-TG levels measured 24 h post-MI were associated with a greater LVEF. In addition, decreased HDL-TG, M-HDL-TG, XL-HDL-TG% and S-HDL-TG measured 24 h post-MI predicted higher ISZ. Categorical analysis of LVEF and ISZ provided similar results with more favorable outcomes for patients with higher HDL-TG levels. No metabolite measure showed a significant interaction with metformin treatment, suggesting that there was no metabolic subgroup of patients in whom metformin was effective.

Our findings suggest beneficial effects of higher triglyceride levels in HDL and in HDL subfractions measured 24 h post-MI on ISZ and LVEF. Clinical studies identified low

admission triglyceride levels as a risk factor for recurrent CV events and mortality in STEMI patients.[22,23] Likewise, low triglyceride levels are associated with a poor prognosis in stroke patients.[24] This contrasts with findings from large-scale case-control and prospective cohort studies indicating that hypertriglycemia is a strong predictor of CV events, even independent of cholesterol levels.[25,26] These epidemiological findings, however, apply to individuals who were not studied during the course of an acute coronary event. Similarly paradoxical findings have been obtained for plasma cholesterol levels. While hypercholesterolemia is an established CV risk factor in the general population, admission LDL-C levels < 70 mg/dl are associated with higher mortality and incidence of heart failure in statin-naïve STEMI patients.[27] The pathogenic mechanisms underlying recurrent CV events shortly after an acute event are still poorly understood. It is possible that in the acute setting HDL-TG plays a distinct role on CV outcome .

VLDL is the most important triglyceride carrier in plasma. The triglyceride content of VLDL showed substantial correlation with HDL-TG 24 h post-MI (Supplemental Figure 1 B). However, only the triglyceride content of very small VLDL particles was associated with LVEF categories. In addition, the TG content of large LDL particles at baseline predicted abnormal LVEF 4 months post-MI. Inhibition of fatty acid uptake by relocation of FAT/CD36 may reduce intracellular fatty acid concentrations[28], resulting in increased extracellular fatty acid levels and diminished lipolysis of lipid-bound triglycerides. This may initially lead to triglyceride enrichment of VLDL and LDL particles, which subsequently transfer excess triglycerides to HDL particles in exchange for cholesteryl esters by the action of cholesteryl ester transfer protein (CETP), thereby increasing the triglyceride content in HDL.[29] In line with this, blood samples of MI patients collected immediately after diagnosis show strong triglyceride enrichment of HDL_2 particles.[30] Higher plasma HDL-TG levels could thus be consequent to inhibition of fatty acid uptake, and coincide with diminished fatty acid oxidation and prevention of further myocardial damage.[31] Larger triglyceride-rich particles are converted to small VLDL subfractions as a result of lipase-mediated delipidation[32], suggesting that triglyceride enrichment may be secondary to initial triglyceride uptake of large VLDL. Larger VLDL particles may be delipidated rapidly, which may explain why the association of triglycerides with LVEF was limited to very small VLDL 24 h post-MI. Similarly, a major proportion of LDL-TG is derived from large VLDL[32], which may partly result from CETP-mediated delipidation of large VLDL. Taken together, early metabolic changes after MI could reflect adaptive mechanisms that promote functional recovery.

While we observed associations of LVEF categories with 24 h post-MI measurements of XS-VLDL-TG% and XS-VLDL-C%, these metabolite measures did not significantly predict LVEF when LVEF was analyzed as a continuous variable. However, the regression model with continuous outcome assumes linearity between metabolite measures and LVEF, whereas categorical analysis of LVEF in combination with multinomial logistic regression

renders the model sensitive to non-linear associations. As shown in Figure 2 A, HDL-TG, M-HDL-TG and S-HDL-TG follow a linear trend across the three LVEF categories, whereas XS-VLDL-TG% and XS-VLDL-C% display non-linear trends.

Limitations

The GIPS-III trial was originally designed to assess differences in LVEF between metformin treated patients controls. We conducted 68 independent tests, raising the possibility that our study was not powered to detect smaller changes. However, we were able to detect a significant effect for alanine levels, which were only slightly increased in the metformin group (median difference: 0.03 mmol/L), demonstrating sufficient power to perform a metabolic profiling analysis. In addition, all patients received intravenous heparin before PCI when baseline blood samples were drawn. Heparin stimulates lipolysis and hence acutely reduces plasma triglyceride levels[33], which is in line with the marked increase in triglyceride levels between baseline and the other time points (Supplemental Table 8). STEMI patients routinely receive heparin before PCI, rendering the results for baseline measurements relevant to clinical settings. These findings measurements should nevertheless be interpreted with caution. Moreover, we performed metabolic profiling in non-fasting blood samples, warranting further research to substantiate our findings under fasting conditions. However, the NMR platform used in our study mainly quantifies lipid measures, which change only slightly after food consumption and show similar associations with cardiovascular risk in fasting and non-fasting individuals.[34]

Conclusions

In summary, our study suggests that metformin treatment started directly after presentation with STEMI produces changes in alanine and XL-HDL-PL% as assessed after 4 months. Higher triglyceride levels in HDL and in HDL subfractions measured 24 h post-MI were predictive of better LVEF and smaller ISZ 4 months post-MI. HDL-TG may thus serve as an early biomarker of left ventricular dysfunction in STEMI patients. However, further studies are required to substantiate the clinical significance of HDL-TG in CV risk prediction and to investigate the biological mechanism underlying associations of metabolic biomarkers with recurrent CV events. Our findings emphasize the utility of high-throughput metabolic profiling as a tool to study drug effects and to identify prognostic biomarkers of LVEF and ISZ.

CLINICAL PERSPECTIVE

Although in-hospital survival after myocardial infarction has improved because of ameliorated treatment strategies, longterm outcome of myocardial infarction is threatened

by the development of heart failure and its associated poor prognosis. Left ventricular ejection fraction and infarct size are key predictors of heart failure development and long-term outcome. However, treatment options for left ventricular dysfunction are limited and the biochemical mechanisms driving functional decline after myocardial infarction are largely unknown. To identify biomarkers predicting post-MI left ventricular ejection fraction and infarct size and to evaluate the metabolic footprint of metformin, we performed metabolic profiling in the GIPS-III cohort (Glycometabolic Intervention as Adjunct to Primary Percutaneous Intervention in ST Elevation Myocardial Infarction), including 233 metabolite measures obtained on admission, 24 hours post-MI and 4 months post-MI. Our results suggest that triglycerides in high-density lipoprotein and high-density lipoprotein subfractions measured 24 hours post-MI predict left ventricular ejection fraction and infarct size. High-density lipoprotein triglyceride may thus serve as an early marker of left ventricular dysfunction after ST-segment–elevated MI. In addition, metformin increased alanine levels and reduced the phospholipid content in very large high-density lipoprotein particles. Our findings emphasize the utility of high-throughput metabolic profiling as a tool to study drug effects and to identify prognostic biomarkers of left ventricular ejection fraction and infarct size.

Supplemental materials are available online.

REFERENCES

1. Sabate M, Cequier A, Iñiguez A, Serra A, Hernandez-Antolin R, Mainar V, et al. Everolimus-eluting stent versus bare-metal stent in ST-segment elevation myocardial infarction (EXAMINATION): 1 year results of a randomised controlled trial. *Lancet*. 2012;380:1482-1490.
2. El Aidi H, Adams A, Moons KG, Den Ruijter HM, Mali WP, Doevendans PA, et al. Cardiac magnetic resonance imaging findings and the risk of CV events in patients with recent myocardial infarction or suspected or known coronary artery disease: a systematic review of prognostic studies. *J Am Coll Cardiol*. 2014;63:1031-1045.
3. Wu E, Ortiz JT, Tejedor P, Lee DC, Bucciarelli-Ducci C, Kansal P, et al. Infarct size by contrast enhanced cardiac magnetic resonance is a stronger predictor of outcomes than left ventricular ejection fraction or end-systolic volume index: prospective cohort study. *Heart*. 2008;94:730-736.
4. Lexis CPH, van der Horst ICC, Lipsic E, van der Harst P, van der Horst-Schrivers ANA, Wolffenbuttel BHR, et al. Metformin in non-diabetic patients presenting with ST elevation myocardial infarction: rationale and design of the glycometabolic intervention as adjunct to primary percutaneous intervention in ST elevation myocardial infarction (GIPS)-III trial. *Cardiovasc Drugs Ther Spons Int Soc Cardiovasc Pharmacother.* 2012;26: 417–426.
5. Lexis CP, van der Horst IC, Lipsic E, Wieringa WG, de Boer RA, van den Heuvel AF, et al. Effect of metformin on left ventricular function after acute myocardial infarction in patients without diabetes: the GIPS-III randomized clinical trial. *JAMA*. 2014;311:1526-1535.
6. Krauss RM, Zhu H, Kaddurah-Daouk R. Pharmacometabolomics of statin response. *Clin Pharmacol Ther*. 2013;94:562-565.
7. Maiso P, Huynh D, Moschetta M, Sacco A, Aljawai Y, Mishima Y, et al. Metabolic signature identifies novel targets for drug resistance in multiple myeloma. *Cancer Res*. 2015;75:2071-2082.
8. Winnike J, Li Z, Wright F, Macdonald J, O'Connell T, Watkins P. Use of pharmaco-metabonomics for early prediction of acetaminophen-induced hepatotoxicity in humans. *Clin Pharmacol Ther*. 2010;88:45-51.
9. Goldberg R, Temprosa M, Otvos J, Brunzell J, Marcovina S, Mather K, et al. Lifestyle and metformin treatment favorably influence lipoprotein subfraction distribution in the Diabetes Prevention Program. *J Clin Endocrinol Metab*. 2013;98:3989-3998.
10. Preiss D, Rankin N, Welsh P, Holman R, Kangas A, Soininen P, et al. Effect of metformin therapy on circulating amino acids in a randomized trial: the CAMERA study. *Diabetic Med*. 2016;33:1569-1574.
11. Eppinga RN, Hartman MH, van Veldhuisen DJ, Lexis CP, Connelly MA, Lipsic E, et al. Effect of Metformin Treatment on Lipoprotein Subfractions in Non-Diabetic Patients with Acute Myocardial Infarction: A Glycometabolic Intervention as Adjunct to Primary Coronary Intervention in ST Elevation Myocardial Infarction (GIPS-III) Trial. *PLOS ONE*. 2016;11:e0145719.
12. Vaarhorst AA, Verhoeven A, Weller CM, Böhringer S, Göraler S, Meissner A, et al. A metabolomic profile is associated with the risk of incident coronary heart disease. *Am Heart J*. 2014;168:45-52.
13. Würtz P, Raiko JR, Magnussen CG, Soininen P, Kangas AJ, Tynkkynen T, et al. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J*. 2012;33:2307-2316.
14. Wang J, Li Z, Chen J, Zhao H, Luo L, Chen C, et al. Metabolomic identification of diagnostic plasma biomarkers in humans with chronic heart failure. *Mol Biosyst*. 2013;9:2618-2626.
15. Rosenson RS. Myocardial injury: the acute phase response and lipoprotein metabolism. *J. Am Coll Cardiol.* 1993;22:933-940.

16. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*. 2015;8:192–206.
17. Stekhoven DJ, Bühlmann P. MissForest – non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28:112-118.
18. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of CV Imaging. *J Am Soc Echocardiogr*. 2015;28:1-39. e14.
19. Komori T, Hotta N, Kobayashi M, Sakakibara F, Koh N, Sakamoto N. Biguanides may produce hypoglycemic action in isolated rat hepatocytes through their effects on L-alanine transport. *Diabetes Re. Clin Pr*. 1993;22:11-17.
20. Song S, Andrikopoulos S, Filippis C, Thorburn AW, Khan D, Proietto J. Mechanism of fat-induced hepatic gluconeogenesis: effect of metformin. *Am J Physiol Endocrinol Metab.* 2001;281:E275-E282.
21. Xu T, Brandmaier S, Messias AC, Herder C, Draisma HH, Demirkan A, et al. Effects of metformin on metabolite profiles and LDL cholesterol in patients with type 2 diabetes. *Diabetes Care*. 2015;38:1858-1867.
22. Cheng Y-T, Liu T-J, Lai H-C, Lee W-L, Ho H-Y, Su C-S, et al. Lower serum triglyceride level is a risk factor for in-hospital and late major adverse events in patients with ST-segment elevation myocardial infarction treated with primary percutaneous coronary intervention- a cohort study. *BMC Cardiovasc. Disord*. 2014;14:143.
23. Khawaja OA, Hatahet H, Cavalcante J, Khanal S, Al-Mallah MH. Low admission triglyceride and mortality in acute coronary syndrome patients. *Cardiol J*. 2011;18:297-303.
24. Weir CJ, Sattar N, Walters MR, Lees KR. Low triglyceride, not low cholesterol concentration, independently predicts poor outcome following acute stroke. *Cerebrovasc Dis*. 2003;16:76-82.
25. Nordestgaard BG, Benn M, Schnohr P, Tybjærg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. *JAMA*. 2007;298:299-308.
26. Sarwar N, Danesh J, Eiriksdottir G, Sigurdsson G, Wareham N, Bingham S, et al. Triglycerides and the risk of coronary heart disease 10 158 incident cases among 262 525 participants in 29 western prospective studies. *Circulation*. 2007;115:450-458.
27. Oduncu V, Erkol A, Kurt M, Tanboğa İH, Karabay CY, Şengül C, et al. The prognostic value of very low admission LDL-cholesterol levels in ST-segment elevation myocardial infarction compared in statin-pretreated and statin-naive patients undergoing primary percutaneous coronary intervention. *Int J Cardiol*. 2013;167:458-463.
28. Heather LC, Pates KM, Atherton HJ, Cole MA, Ball DR, Evans RD, et al. Differential translocation of the fatty acid transporter, FAT/CD36, and the glucose transporter, GLUT4, coordinates changes in cardiac substrate metabolism during ischemia and reperfusion. *Circ Heart Fail*. 2013;6:1058-1066.
29. Kappelle PJ, van Tol A, Wolffenbuttel BH, Dullaart RP. Cholesteryl ester transfer protein inhibition in CV risk management: ongoing trials will end the confusion. *Cardiovasc Ther.* 2011;29:e89-99.
30. Cho K-H, Shin D-G, Baek S-H, Kim J-R. Myocardial infarction patients show altered lipoprotein properties and functions when compared with stable angina pectoris patients. *Exp Mol Med*. 2009;41:67-76.

31. Liu Q, Docherty JC, Rendell JC, Clanachan AS, Lopaschuk GD. High levels of fatty acids delay the recovery of intracellular pH and cardiac efficiency in post-ischemic hearts by inhibiting glucose oxidation. *J Am Coll Cardiol*. 2002;39:718-725.
32. Packard CJ and Shepherd J. Lipoprotein heterogeneity and apolipoprotein B metabolism. *Arterioscler Thromb Vasc Biol.* 1997;17:3542-3556.
33. Brunner MP, Shah SH, Craig DM, Stevens RD, Muehlbauer MJ, Bain JR, et al. Effect of heparin administration on metabolomic profiles in samples obtained during cardiac catheterization. *Circ Cardiovasc Genet*. 2011;4:695-700.
34. Nordestgaard, BG, Langsted A, Mora S, Kolovou G, Baum H, Bruckert E, et al. Fasting is not routinely required for determination of a lipid profile: clinical and laboratory implications including flagging at desirable concentration cut-points – a joint consensus statement from the European Atherosclerosis Society and European Federation of Clinical Chemistry and Laboratory Medicine. *Eur Heart J.* 2016;37:1944-1958.

Chapter 7

Statin-effects on Metabolic Profiles: Data from the Prevend It Trial

Ruben N. Eppinga*, Daniel Kofink*, Wiek H. van Gilst, Stephan J.L. Bakker, Robin P.F. Dullaart, Pim van der Harst†, Folkert W. Asselbergs†

* both authors contributed equally, † both authors contributed equally

Adapted from Circ Cardiovasc Genet. 2017 Dec;10(6). pii: e001759

ABSTRACT

Background: Statins lower cholesterol by inhibiting HMG-CoA reductase, the rate- limiting enzyme of the metabolic pathway that produces cholesterol and other isoprenoids. Surprisingly little is known about their effects on metabolite and lipoprotein subclass profiles. We therefore investigated the molecular changes associated with pravastatin treatment compared to placebo administration, using a nuclear magnetic resonance (NMR)-based metabolomics platform.

Methods and results: We performed metabolic profiling of 231 lipoprotein and metabolite measures in the PREVEND IT study, a placebo-controlled randomized clinical trial designed to test the effects of pravastatin (40 mg once daily) on cardiovascular risk. Metabolic profiles were assessed at baseline and after 3 months of treatment. Pravastatin lowered low-density lipoprotein cholesterol (LDL-C; change in SD units [95% CI]: -1.01 [-1.14, -0.88]), remnant cholesterol (change in SD units [95% CI]: -1.03 [95% CI : -1.17, -0.89]) and apolipoprotein B (apoB, change in SD units [95% CI]: -0.98 [95% CI : -1.11, -0.86]) with similar effect magnitudes. In addition, pravastatin globally lowered levels of lipoprotein subclasses, with the exception of high-density lipoprotein (HDL) subclasses, which displayed a more heterogeneous response pattern. The lipid lowering effect of pravastatin was accompanied by selective changes in lipid composition, particularly in the cholesterol content of very low-density lipoprotein (VLDL) particles. In addition, pravastatin reduced levels of several fatty acids, but had limited effects on fatty acid ratios.

Conclusions: These randomized clinical trial data demonstrate the widespread effects of pravastatin treatment on lipoprotein subclass profiles and fatty acids.

INTRODUCTION

Statins hamper cholesterol production in the liver through inhibition of HMG-CoA reductase, which, in turn, stimulates hepatic synthesis of low-density lipoprotein (LDL) receptors as a compensatory mechanism. These receptors bind to apoB-rich lipoproteins and facilitate their absorption by hepatocytes, leading to a further reduction in plasma cholesterol levels.[1] The cardiovascular risk reduction achieved through statins is believed to primarily result from their LDL cholesterol (LDL-C) lowering properties.[2] Lowering of LDL-C has therefore been identified as the primary treatment target of statin therapy.[3] However, statins act early in the mevalonate pathway and have the potential to extensively modify the metabolic profile in addition to their effect on cholesterol metabolism. This has led to the hypothesis that statins may provide cardioprotective benefits beyond LDL-C reduction. While there is mounting evidence underpinning the therapeutic capacities of such pleiotropic statin effects,[4-6] little is known about the underlying molecular pathways.

Nuclear magnetic resonance (NMR)-based metabolic profiling has evolved into a versatile high-throughput tool for biomarker discovery that allows simultaneous quantifications of numerous molecules, ranging from amino acids to a variety of lipoprotein subclass measures. Metabolic profiling has been widely used both in epidemiology and in drug research.[7-9] Better characterization of the metabolic footprint of statins may provide novel insights into their mechanisms of action and help guide drug discovery. A recent study of four observational population-based cohorts investigated the longitudinal effects of statins on metabolic profiles by comparing users to non-users, followed by confirmatory Mendelian randomization analysis.[9] Besides cholesterol lowering, statins influenced fatty acid levels, whereas amino acids and other metabolites were not substantially altered. While this study revealed extensive changes in routine lipid measures, little is known about the effect of statins on lipoprotein subclass profiles, even though mounting evidence suggests distinct roles for lipoprotein subclasses in the pathophysiology of cardiovascular disease.[10-12] In addition, no study has yet comprehensively investigated the metabolic effects of statin therapy in a placebo-controlled randomized setting. Here we present the first data on pravastatin treatment derived from the Prevention of Renal and Vascular End-stage Disease Intervention Trial (PREVEND IT) study, a randomized placebo-controlled clinical trial. In addition to previously quantified parameters, including lipids, fatty acids, amino acids and glycolysis metabolites, we report results for over 160 measures of lipoprotein subclasses. An overview of lipoprotein subclasses is given in Table 1.

Table 1. Average particle size of lipoprotein subclasses

Lipoprotein	Subclass	Average particle diameter (in nm)*
VLDL	XXL	>75
	XL	64.0
	L	53.6
	M	44.5
	S	36.8
	XS	31.3
IDL		28.6
LDL	L	25.5
	M	23.0
	S	18.7
HDL	XL	14.3
	L	12.1
	M	10.9
	S	8.7

Average particle size (diameter in nm) for different lipoprotein subclasses. *Average particle diameters adapted from [8]. Cut points for size ranges can be approximated by the midpoint between the average diameters of two consecutive lipoprotein subclasses, e.g. the lower bound of XS-VLDL is approximately 30 nm. HDL: high-density lipoprotein; IDL: intermediate-density lipoprotein; LDL: low-density lipoprotein; VLDL: very low-density lipoprotein; XXL: extremely large; XL: very large; L: large; M: medium; S: small; XS: very small.

MATERIALS AND METHODS

Subjects

Details on the PREVEND IT study have been published elsewhere.[13] Briefly, PREVEND IT is a double-blind, placebo-controlled clinical trial, in which participants were randomized to 20 mg fosinopril or matching placebo and 40 mg pravastatin or matching placebo. PREVEND IT participants were recruited from the PREVEND program, which investigated the influence of microalbuminuria on cardiovascular and renal risk. The main inclusion criteria for PREVEND IT were a urine albumin concentration of >10 mg/L in one morning spot sample and at least once a concentration of 15 to 300 mg/24 h in two successive 24-hour urine samples, a blood pressure of <160/100 mm Hg, no hypertensive treatment and a total serum cholesterol concentration < 8.0 mmol/L (or <5.0 mmol/L in case of prior myocardial infarction) and no lipid-lowering treatment. 864 subjects were randomized to receive study medication (see above) after giving informed consent. Blood samples for metabolic profiling were limited by sample availability and could be obtained in 394 participants at baseline and after 3 months of treatment. The study was approved by the Institutional Review Board and was conducted in according to the guidelines of the declaration of Helsinki.

Laboratory measurements

Fasting blood samples were drawn before treatment onset (baseline) and at the 3-month medical review (N=394). Metabolic profiling was performed in EDTA anticoagulated plasma samples using high-throughput 1H NMR metabolomics (Brainshake Ltd, Helsinki, Finland), as previously described[7]. This method provides accurate quantification of 231 lipoprotein and metabolite measures, including routine lipids, lipoprotein profiles with 14 lipoprotein subfractions, glycolysis related metabolites, amino acids, ketone bodies, fluid balance related metabolites and one inflammatory marker (Supplemental Table 1). Recent studies have demonstrated that NMR measurements quantified with this platform are in good agreement with routine clinical chemistry assays.[8] Representative coefficients of variation (CVs) for this platform have been reported elsewhere.[14]

Statistical analysis

Correlations between different lipoprotein and metabolite measures were calculated using Spearman's correlation coefficients. The effect of statin treatment on each NMR measure was assessed by linear regression on the change during the treatment period, as previously described.[9] The effect estimate (regression coefficient) of this regression model can be interpreted as the longitudinal change of a NMR measure attributable to pravastatin treatment. To facilitate comparison between different lipoprotein and metabolite measures, differences between pre- and post-treatment values were scaled to baseline SD units. Consequently, statin effects on NMR measures are expressed in baseline SD units. We additionally performed a sensitivity analysis adjusted for sex as the pravastatin group showed a higher percentage of male patients. Since many NMR measures were highly correlated (see Supplemental Table 1), we accounted for multiple testing by correcting the nominal level of significance for the number of independent tests, which was estimated by the method of Li and Ji,[15] using the matrix spectral decomposition (matSpD) tool (http://gump.qimr.edu.au/general/ daleN/matSpD/). The number of independent tests was estimated to be 85, yielding a corrected significance threshold of 0.05/85=0.00059.

RESULTS

Baseline characteristics and NMR measures

Baseline characteristics of all patients included in this study are listed in Table 2. Of 394 participants, 195 received pravastatin and 199 placebo during the 3-month treatment period. A summary of all 231 lipoprotein and metabolite measures can be found in Supplemental Table 1. NMR and available clinical chemistry measures showed strong correlations for baseline and post-treatment measurements (Supplemental Table 2),

Table 2. Baseline characteristics

Variable	Placebo (n=199)	Pravastatin (n=195)
Age (years)	50.6±11.1	51.5±11.5
Male	121 (60.8)	141 (72.3)
BMI (kg/m^2)	26.5±4.5	26.3±4.1
Current smoker	79 (39.7)	82 (42.1)
SBP (mm Hg)	130.6±17.3	131.6±18.3
DBP (mm Hg)	75.8±9.9	76.6±9.4
Cholesterol (mmol/l)	5.9±1.0	5.9±1.1
HDL (mmol/l)	1.0±0.3	1.0±0.3
LDL (mmol/l)	4.1±0.9	4.2±1.0
Triglycerides (mmol/l)	1.3 (0.9-1.9)	1.4 (0.9-1.9)
Glucose (mmol/l)	4.9 (4.5-5.3)	4.9 (4.5-5.3)
Creatinine (µmol/l)	84.0±15.1	86.4±13.2
Medication use		
Beta-blockers	4 (2.0)	0 (0.0)
Nitrate	2 (1.0)	0 (0.0)
Diuretics	4 (2.0)	0 (0.0)
Calcium channel blockers	0 (0.0)	1 (0.5)
Digoxin	1 (0.5)	2 (1.0)

Discrete variables are expressed as absolute count (%) and continuous variables as mean±SD or median (interquartile range). Lipids, glucose and creatinine as measured by clinical chemistry. SD: standard deviation; BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; HDL: high-density lipoprotein; LDL: low-density lipoprotein; SD: standard deviation.

indicating consistency between different analytical methods. Heat maps of correlations between NMR measures are displayed in Supplemental Figure 1, revealing substantial correlation within lipoprotein subclasses, between amino acids and between fatty acids.

Statin effects

We compared longitudinal changes of NMR measures between the pravastatin group and controls, using linear regression. To facilitate comparison between different measures, differences between pre- and post-treatment values were scaled to baseline SD units. After the 3-month treatment period, a total of 150 NMR measures were significantly altered ($p<0.00059$) between the pravastatin group and the control group. Absolute concentration changes are given for all lipoprotein and metabolite measures in Supplemental Table 3. Additional sensitivity analysis adjusted for sex provided similar findings, suggesting that our results were not confounded by the imbalance in sex ratio between the pravastatin group and the control group (Supplemental Table 4).

As compared with placebo, pravastatin reduced levels of conventional lipid measures (Figure 1), including total serum cholesterol (change associated with pravastatin in SD units [95% CI]: -1.01 [-1.14, -0.88]; $p=7.3x10^{-41}$), LDL-C (change in SD units [95% CI]: -1.01 [-1.13, -0.88]; $p=6.7x10^{-42}$) and total serum triglycerides (change in SD units [95% CI]: -0.46 [-0.60, -0.33]; $p=1.8x10^{-11}$), whereas HDL-C levels were not affected by statin treatment (change in SD units [95% CI]: -0.01 [-0.11, 0.09]; p=0.829). However, pravastatin significantly increased cholesterol in large lipid-rich HDL2 particles (change in SD units [95% CI]: 0.18 [0.08, 0.27]; p=0.00048) and decreased cholesterol in small less dense HDL3 particles (change in SD units [95% CI]: -0.69 [-0.87, -0.51]; $p= 3.1x10^{-13}$).

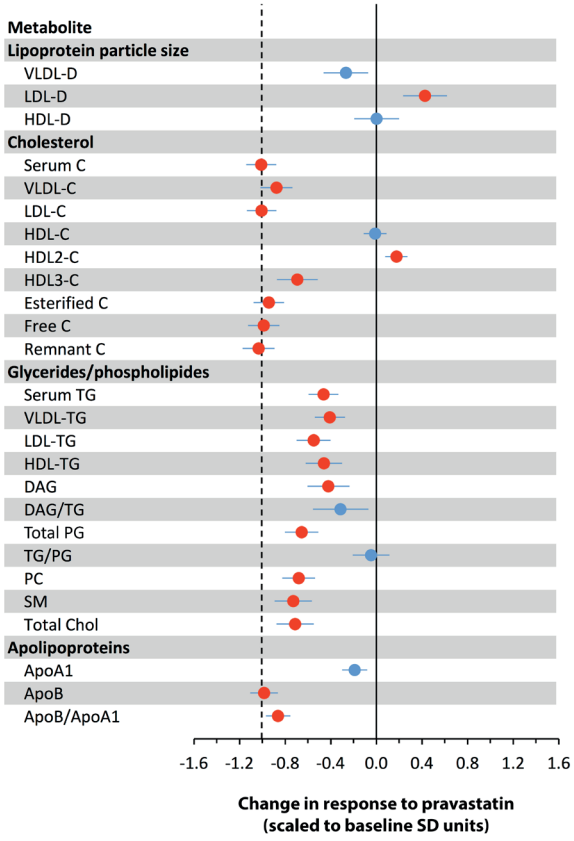


Figure 1. Lipids and lipid-related NMR measures

Concentration changes in lipids and lipid-related measures associated with pravastatin treatment (n=195) compared with placebo treatment (n=199). Effect estimates indicate changes over the treatment period (3 months) associated with pravastatin treatment in baseline SD-units. Error bars represent 95% confidence intervals. The dotted line shows the effect estimate for LDL-C. Red marks indicate significant changes (p<0.00059). HDL: high-density lipoprotein; LDL: lowdensity lipoprotein; VLDL: very low-density lipoprotein; D: diameter; C: cholesterol; TG: triglycerides; DAG: diacylglycerol; PG: phosphoglycerides; PC: phosphatidylcholine; SM: sphingomyelins; Total Chol: total cholines; apoA1: apolipoprotein A1; apoB: apolipoprotein B.

Moreover, pravastatin treatment markedly lowered remnant cholesterol levels (change in SD units [95% CI]: -1.03 [-1.17, -0.89]; $p=2.0x10^{-38}$), which reflects the total cholesterol content in very large-density lipoprotein (VLDL; change in SD units [95% CI]: -0.88 [-1.02, -0.74]; $p=2.1x10^{-29}$) and intermediate-density lipoprotein (IDL; change in SD units [95% CI]: 1.03 [-1.16, -0.89]; $p=1.3x10^{-39}$). The effect of pravastatin on apolipoprotein B (apoB; change in SD units [95% CI]: -0.98 [-1.11, -0.86]; $p=1.1x10^{-44}$) was comparable to the change in LDL-C. Pravastatin globally lowered levels of VLDL, LDL and IDL subclasses (Figure 2), whereas changes in HDL subclasses were less consistent, with significant increases across large HDL subclasses measures and a reduction in small and very large HDL-C.

Particle concentrations of all VLDL, IDL and LDL subclasses decreased in response to statin treatment. IDL was the subclass with the greatest change in particle concentration (change in SD units [95% CI]: -1.04 [95% CI: -1.17 to -0.91]; $p=7.6x10^{-45}$). In addition, we analyzed the lipid composition of different lipoprotein subclasses, expressed as the ratio of individual lipid concentrations to the total lipid concentration (Figure 3). Pravastatin treatment markedly lowered the cholesterol and cholesteryl ester to total lipids ratio in IDL and across all LDL subclasses, concomitant with an elevated relative content of free cholesterol and phospholipids in LDL. Furthermore, pravastatin selectively reduced cholesterol ratios in small and medium VLDL particles. In parallel with cholesterol and triglycerides, pravastatin lowered fatty acid concentrations (Figure 4), particularly ω-6 fatty acids (change in SD units [95% CI]: -0.85 [95% CI : -1.00, -0.71]; $p=3.5x10^{-26}$), total polyunsaturated fatty acids (PUFA, change in SD units [95% CI]: -0.84 [95% CI : -0.98, -0.69]; $p= 3.4x10^{-26}$). By contrast, pravastatin treatment only altered the saturated fatty acid to total fatty acid ratio (SFA/FA; change in SD units [95% CI]: 0.51 [95% CI : 0.29, 0.74]; $p= 9.4x10^{-6}$) and the lineolic acid to total fatty acid ratio (LA/FA; change in SD units [95% CI]: -0.35 [95% CI : 0.48, 0.21]; $p= 7.2x10^{-7}$), but produced no changes in other fatty acid ratios. Glycolysis-related metabolites, amino acids and other metabolites remained unchanged.

We next evaluated the effect of pravastatin on correlations between different NMR measures. Results are illustrated in a correlation difference map that provides a post-treatment comparison between the pravastatin group and controls (Supplemental Figure 2). Pravastatin induced negative associations between the relative cholesterol content of medium HDL and cholesterol levels in small VLDL, IDL and LDL. We observed similar, but weaker effects for absolute cholesterol concentrations in medium HDL. Conversely, pravastatin strengthened or induced positive correlations between the phospholipid-to-total lipids ratio in medium HDL and lipid concentrations in other lipoproteins. Furthermore, correlations between absolute lipid concentrations and the relative lipid content were altered across VLDL subclasses. Finally, lactate and pyruvate showed weaker associations with lipid concentrations in VLDL following pravastatin treatment.

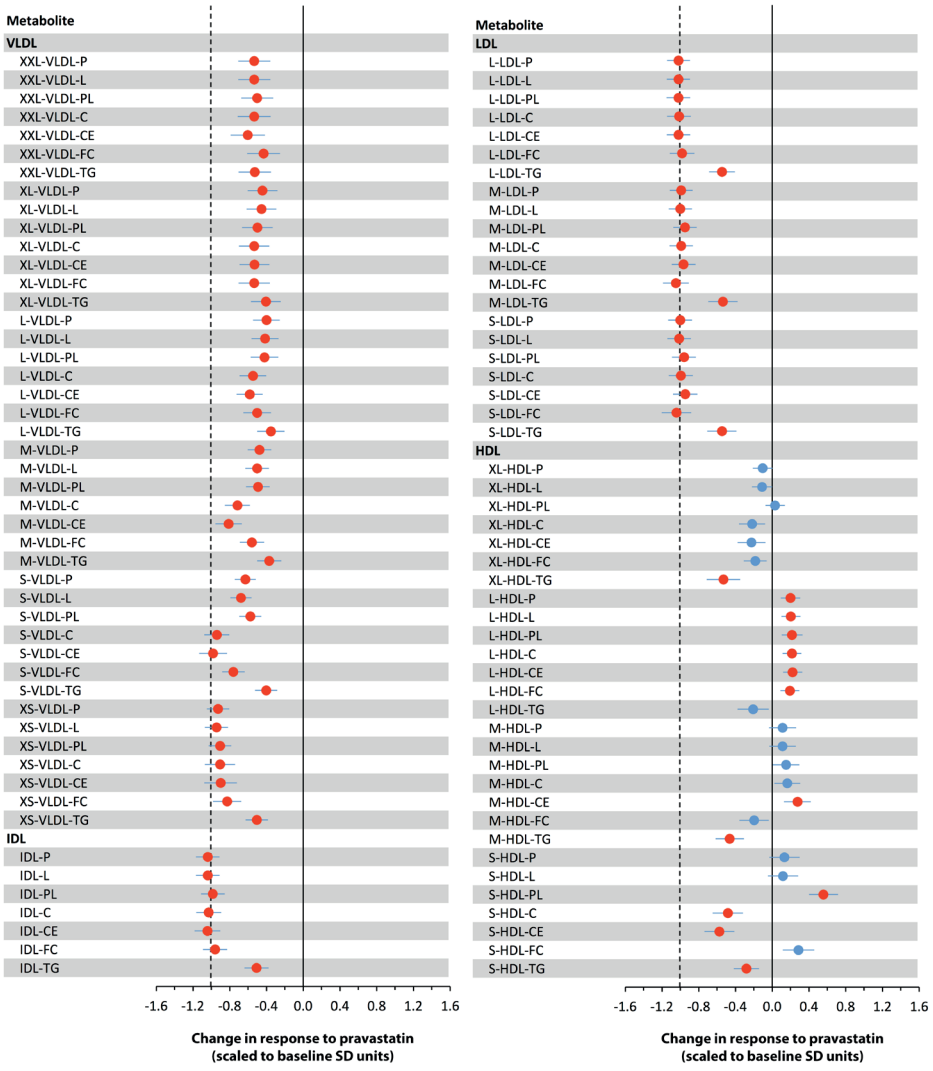


Figure 2. Lipid concentrations in different lipoprotein subclasses

Changes in lipid concentrations across lipoprotein subclasses associated with pravastatin treatment (n=195) compared with placebo treatment (n=199). Effect estimates indicate changes over the treatment period (3 months) associated with pravastatin treatment in SD-units. Error bars represent 95% confidence intervals. The dotted line shows the effect estimate for LDL-C. Red marks indicate significant changes ($p<0.00059$). XXL: extremely large; XL: very large; L: large; M: medium; S: small; XS: very small; HDL: high-density lipoprotein; IDL: intermediate-density lipoprotein; LDL: low-density lipoprotein; VLDL: very low-density lipoprotein; P: particle concentration; L: total lipids; PL: phospholipids; C: cholesterol; CE: cholesteryl esters; FC: free cholesterol; TG: triglycerides.

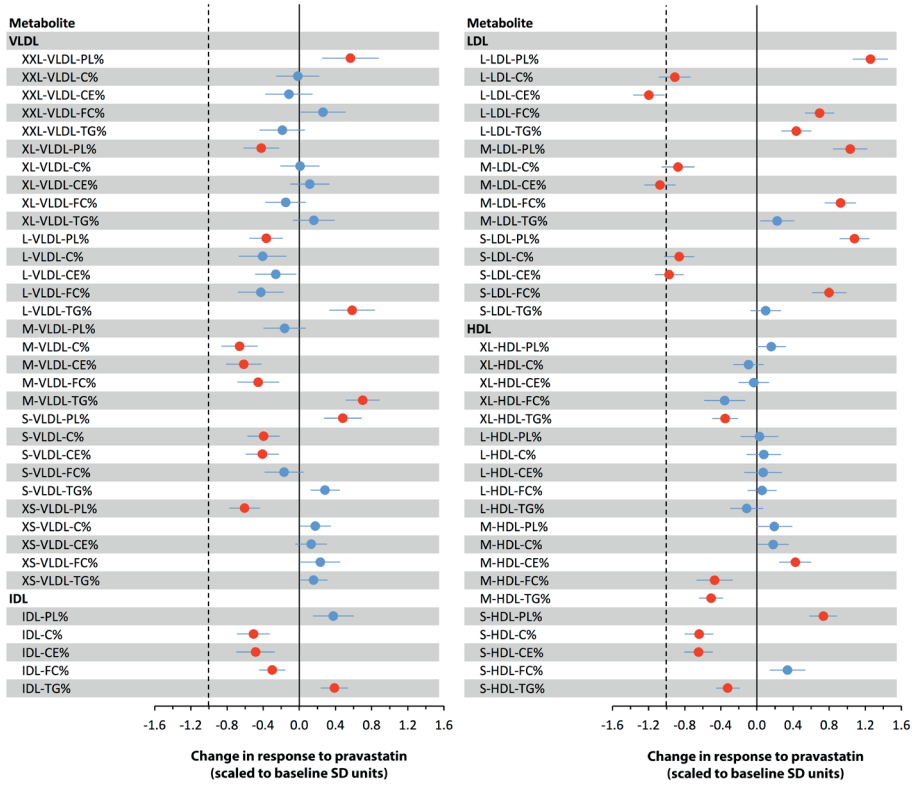


Figure 3. Lipid composition of lipoprotein subclasses

Changes in lipid composition of lipoprotein subclasses associated with pravastatin treatment (n=195) compared with placebo treatment (n=199). Effect estimates indicate changes over the treatment period (3 months) associated with pravastatin treatment in baseline SD-units. Error bars represent 95% confidence intervals. The dotted line shows the effect estimate for LDL-C. Red marks indicate significant changes ($p<0.00059$). %: lipid concentration relative to total lipid concentration; XXL: extremely large; XL: very large; L: large; M: medium; S: small; XS: very small; HDL: high-density lipoprotein; IDL: intermediate-density lipoprotein; LDL: low-density lipoprotein; VLDL: very low-density lipoprotein; P: particle concentration; L: total lipids; PL: phospholipids; C: cholesterol; CE: cholesteryl esters; FC: free cholesterol; TG: triglycerides.

DISCUSSION

This is the first placebo-controlled NMR study to assess metabolic changes associated with statin treatment, using data from the PREVEND IT trial. Our study adds to previous findings from observational NMR studies and additionally explored statin-induced changes in over 160 novel measures of lipid concentrations and lipid composition for 14 lipoprotein subclasses. Besides the well-known effects on LDL-C, statins altered a wide range of lipids and concentrations of fatty acids. These findings are supported by observational studies comparing statin users to non-users,[9,16] and fit with previous clinical

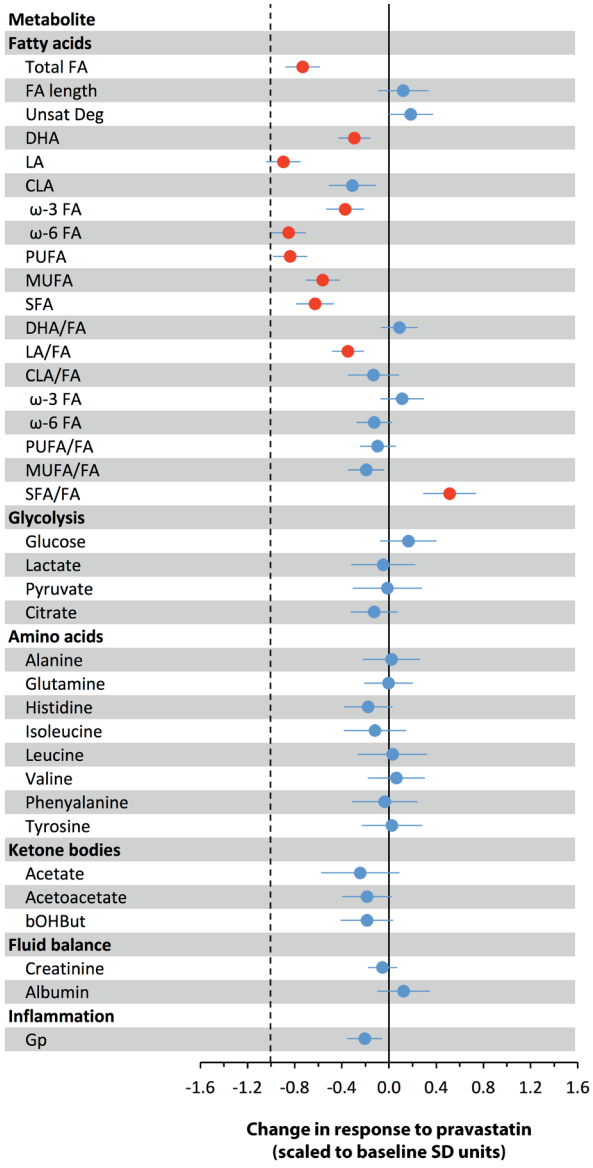


Figure 4. Fatty acids, amino acids and other metabolites

Concentration changes in fatty acids, amino acids and other metabolites associated with pravastatin treatment (n=195) compared with placebo treatment (n=199). Effect estimates indicate changes over the treatment period (3 months) associated with pravastatin treatment in baseline SD-units. Error bars represent 95% confidence intervals. The dotted line shows the effect estimate for LDL-C. Red marks indicate significant changes ($p<0.00059$). FA: fatty acids; Unsat Deg: degree of unsaturation; DHA: docosahexaenoic acid; LA: linoleic acid; CLA: conjugated linoleic acid; ω-3 FA: omega-3 fatty acids; ω-6 FA: omega-6 fatty acids; PUFA: polyunsaturated fatty acids; MUFA: monounsaturated fatty acids; SFA: saturated fatty acids; bOHbut: 3-hydroxybutyrate; Gp: glycoprotein acetyls.

trial data on fatty acids.[17] By contrast, pravastatin treatment only altered LA/FA and SFA/FA, but had no effect on other fatty acid ratios. In addition, pravastatin globally lowered levels of lipoprotein subclasses, except for HDL concentrations, which displayed a more intricate response pattern. Detailed lipid profiling revealed that the substantial lowering of VLDL-C, IDL-C and LDL-C was paralleled by more selective changes in lipid composition of different lipoprotein particles. Finally amino acids and other metabolites were not affected by statin treatment.

Statins not only act on LDL, but also on other apoB-rich lipoproteins. In our study, pravastatin reduced apoB and LDL-C with similar effect magnitudes. ApoB has been proposed as a more robust cardiovascular risk marker than LDL-C, supporting the use of apoB as an alternative treatment target for statin therapy.[18,19]

Consistent with previous findings,[9,20] statin treatment substantially lowered cholesterol in apoB-containing, triglyceride-rich remnant particles, including IDL and VLDL. Since VLDL is the main carrier of triglycerides, remnant cholesterol is strongly associated with triglyceride levels. Although triglycerides are well-established markers of cardiovascular risk, their relationship with atherogenesis is not straightforward.[21] By contrast, remnant cholesterol is likely to play a causal role in cardiovascular disease risk.[22,23] In line with this, remnant cholesterol is associated with both ischemic heart disease and low-grade inflammation.[24] Compared with VLDL-C and IDL-C, HDL-C showed a more complex response to statin treatment, with cholesterol depletion of small HDL3 particles and slight cholesterol enrichment of larger HDL2 particles. Prospective cohort studies have consistently reported inverse associations between HDL-C levels and risk of cardiovascular disease,[25] whereas findings from recent Mendelian randomization studies[26,27] and the failure of HDL-raising drugs to improve cardiovascular outcomes[28] may argue against a causal role for HDL-C in cardiovascular disease per se. Findings from experimental studies suggest that HDL3 and HDL2 differ in their cardioprotective capacities.[29] However, the relationship between different HDL subclasses and cardiovascular risk remains a matter of debate as results from observational studies are inconclusive.[30]

It has been suggested that small dense LDL particles are more atherogenic than larger LDL species as they are readily taken up by the arterial wall, are cleared from circulation at reduced rates due to their low affinity for LDL receptors and are more susceptible to oxidation, promoting the formation of atherosclerotic plaques.[11,13] This is supported by large cohort studies, demonstrating that concentrations of small rather than large LDL particles are associated with future cardiovascular risk after adjustment for non-lipid risk factors.[31-33] However, effect estimates for small LDL particles are not superior to total LDL concentrations and do not improve risk prediction beyond routine lipid measures.[31] Moreover, a systematic review of NMR studies found no association of LDL subclasses with cardiovascular disease after adjustment for other lipid measurements.[34] Experimental findings suggest that, similar to LDL particles, the atherogenic capacity

of VLDL may depend on particle size as large VLDL subpopulations are unable to enter the arterial wall and are thus less likely to contribute to the formation of atherosclerotic plaques.[35] However, there is little evidence from clinical studies that smaller and larger VLDL particles differ in their atherogenic potential. While different lipoprotein subclasses may play distinct roles in the pathophysiology of cardiovascular disease, pravastatin treatment lowered lipoprotein particle concentrations and lipid concentrations across VLDL, IDL and LDL subclasses, which may be an indirect consequence of enhanced clearance and/or reduced synthesis of these lipoproteins.

While the cholesterol-to-total-lipids ratio was decreased in IDL and all LDL subpopulation, pravastatin selectively reduced the cholesterol content of small and medium VLDL, raising the possibility that statins specifically target potentially atherogenic VLDL subpopulations.[35] At the same time, pravastatin lowered the triglyceride to total lipids ratio across all HDL subpopulations, but increased the triglyceride content of several VLDL and LDL subclasses as well as IDL. These changes in lipid composition may be attributable to statin effects on the reverse cholesterol transport pathway, in which cholesteryl ester transfer protein (CETP) transfers cholesteryl esters from HDL to triglyceride-rich, apoB-containing lipoproteins (LDL, IDL and VLDL) in exchange for triglycerides.[36] The statin-induced decrease in lipoprotein concentrations is associated with reduced CETP activity, resulting in cholesterol enrichment of HDL and cholesterol depletion of apoB containing lipoproteins.[37-39] Consistent with reduced CETP activity, pravastatin induced negative correlations between the cholesterol content of medium HDL and cholesterol levels in non-HDL particles. Correlation coefficients for other HDL subpopulations, however, were only moderately altered after pravastatin treatment.

The relative reduction in LDL cholesterol was associated with no or only minor changes in the triglyceride content of LDL particles. By contrast, there was triglyceride enrichment of IDL as well as medium and large VLDL particles. Statin-induced lowering CETP activity may also hamper TG transfer from VLDL and IDL to LDL,[40] which would explain the increased relative triglyceride content of IDL and VLDL. Besides lowering the cholesterol content of IDL and LDL, pravastatin treatment led to a relative increase in phospholipids and free cholesterol, which may result from reduced enzymatic cholesterol esterification due to blocked cholesterol synthesis.[41] Taken together, detailed analysis of lipoprotein subclasses revealed selective changes in lipid composition, whereas lipid concentrations were reduced across all VLDL, IDL and LDL subclasses, following pravastatin treatment.

Several studies have shown that besides lipid lowering, statins alter fatty acid levels.[9,16,17] Since the vast majority of circulating fatty acids are bound in triglycerides, cholesteryl esters and phospholipids,[17] the reduction in fatty acid levels associated may result from the statin-induced decrease in lipoproteins providing the main source of circulating lipids. Alternatively, statins may interfere with fatty acid metabolism through different molecular pathways. Simvastatin treatment increases metabolic

indices indicating elevated activity of elongases and desaturases,[17] two enzymes that catalyze the formation of highly unsaturated long-chain fatty acids. Moreover, statin treatment may stimulate hepatic uptake and beta-oxidation of fatty acids by enhancing expression of peroxisome proliferator-activated receptors (PPARs).[42] We observed elevated SFA/FA and reduced LA/FA, but no effects on other fatty acid ratios, which more appropriately reflect fatty acid metabolism than fatty acid concentrations given the lipoprotein-lowering effect of statins. By contrast, a recent observational study reported stronger effects on docosahexaenoic acid (DHA)/FA, whereas SFA/FA was unchanged after statin treatment.[9] In this study, however, information on statin type and dosage was not available. Consistent with our findings, data from a clinical trial suggest that simvastatin does not enhance DHA/FA.[17] Interestingly, studies comparing different statins reported that pravastatin, in contrast to other statins, did not influence selected fatty acid ratios, indicating that changes in fatty acid metabolisms depend on the statin type.[43,44] While the decrease in LA/FA is supported by other studies,[9,17] the underlying metabolic processes remain unclear. Statins increase lecithin:cholesterol acyltransferase (LCAT) activity, which synthesizes cholesteryl esters from cholesterol and fatty acids.[45] Since LA is the preferential substrate of LCAT, elevated LCAT activity would be consistent with higher LA/FA. Collectively, changes in absolute fatty acid levels are mainly driven by statin-induced lipid lowering, whereas statin effects on fatty acid metabolism remain uncertain and may differ between statins.

Pyruvate and lactate showed weaker correlations with VLDL-related measures after pravastatin treatment, whereas absolute concentrations of these two metabolites remained unchanged. In addition to producing lactate, pyruvate is involved in glucose and fatty acid metabolism by forming acetyl-coenzyme A, which is involved in fatty acid synthesis.[46] Fatty acids, in turn, are joined with glycerol to form triglycerides, the main component of VLDL. Pyruvate and lactate as a metabolic product of pyruvate are thus associated with enhanced hepatic VLDL synthesis and consequently should show a positive correlation with serum VLDL levels. This is in line with the correlation patterns of pyruvate and lactate in the placebo group (Supplemental Figure 2B). Statins, however, facilitate hepatic uptake of non-HDL particles, including VLDL, by increasing LDL receptor activity.[1] The resulting decrease in VLDL levels coupled with unchanged pyruvate and lactate levels is consistent with weaker correlations in the pravastatin group (Supplemental Figure 2A).

Our study was powered to detect a large number of significant changes in lipoprotein and metabolite measures after pravastatin treatment, underscoring the strengths of a placebo controlled randomized setting with pre/post treatment comparisons, which limits potential sources of confounding to a minimum. We report associations for 231 NMR measures, including over 160 novel measures of lipid concentrations and lipid composition for different lipoprotein subclasses. No other study has assessed the

effect of statins on lipoprotein subclasses in such detail. However, further research is warranted to confirm our findings on lipoprotein subclasses as we did not replicate our results in an independent study. In comparison with a recent observational study that used the same NMR metabolomics platform,[9] we observed more moderate effects of statin treatment on several lipid measures, including LDL-C, apoB and apoA1. Würtz et al. compared statin users, who commenced statin treatment, to non-users. While information on statin type and dosage was not available for this study, all statin users had an indication for statin therapy, such as hypercholesterolemia, suggesting that many of them underwent aggressive treatment. In our study, however, participants were randomly assigned to a moderate dose of a relatively week statin,[47] which may account for the lower effect estimates.

In conclusion, metabolic profiling in a randomized clinical trial revealed causal associations of statin treatment with globally reduced lipid levels across lipoprotein subclasses, accompanied by more selective changes in the lipid composition of lipoproteins. Additionally, pravastatin treatment lowered fatty acid concentrations, but had limited effects on fatty acid ratios. In line with previous findings[9], statin treatment did not alter concentrations of non-lipid measures, such as amino acids and glycolysis-related metabolites, suggesting that these metabolites do not reflect pleiotropic statin effects. Our findings demonstrate that high-throughput metabolic profiling is emerging as a powerful tool to dissect a drug's metabolic footprint, providing important information that may be used to improve current treatments.

Supplemental materials are available online.

REFERENCES

1. Brautbar A, Ballantyne CM. Pharmacological strategies for lowering LDL cholesterol: statins and beyond. *Nat Rev Cardiol.* 2011;8:253-265.
2. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet.* 2010;376:1670-1681.
3. Catapano AL, Graham I, De Backer G, Wiklund O, Chapman MJ, Drexel H, et al. 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *Eur Heart J.* 2016;37:2999-3058.
4. Fichtlscherer S, Schmidt-Luckke C, Bojunga S, Rössig L, Heeschen C, Dimmeler S, et al. Differential effects of short-term lipid lowering with ezetimibe and statins on endothelial function in patients with CAD: clinical evidence for 'pleiotropic' functions of statin therapy. *Eur Heart J.* 2006; 27:1182-1190.
5. Liu P-Y, Liu Y-W, Lin L-J, Chen J-H, Liao JK. Evidence for statin pleiotropy in humans: differential effects of statins and ezetimibe on rho-associated coiled-coil containing protein kinase activity, endothelial function, and inflammation. *Circulation. 2009*;119:131-138.
6. McMurray JJ, Kjekshus J, Gullestad L, Dunselman P, Hjalmarson A, Wedel H, et al. Effects of statin therapy according to plasma high-sensitivity c-reactive protein concentration in the controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA) a retrospective analysis. *Circulation.* 2009;120:2188-2196.
7. Eppinga RNE, Kofink D, Dullaart RPF, Dalmeijer GW, Lipsic E, van Veldhuisen DJ, et al. Effect of metformin on metabolites and relation with myocardial infarct size and left ventricular ejection fraction after myocardial infarction. *Circ Cardiovasc Genet.* 2017;10:e001564.
8. Würtz P, Kangas AJ, Soininen P, Lawlor, DA, Smith GD, Ala-Korpela M. Quantitative Serum NMR Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technology. *Am J Epidemiol.* 2017
9. Würtz P, Wang Q, Soininen P, Kangas AJ, Fatemifar G, Tynkkynen T, et al. Metabolomic profiling of statin use and genetic inhibition of HMG-CoA reductase. *J Am Coll Cardiol.* 2016;67:1200-1210.
10. Berneis KK, Krauss, RM. Metabolic origins and clinical significance of LDL heterogeneity. J Lipid Res. 2002;43:1363-1379.
11. Camont L, Chapman MJ, Kontush A. Biological activities of HDL subpopulations and their relevance to cardiovascular disease. Trends Mol Med. 2011;17:594-603.
12. Kwiterovich P. Clinical relevance of the biochemical, metabolic, and genetic factors that influence low-density lipoprotein heterogeneity. Am J Cardiol 2002;90:30I-47I
13. Asselbergs FW, Diercks G, Hillege H, van Boven AJ, Janssen WM, Voors AA, et al. Effects of fosinopril and pravastatin on cardiovascular events in subjects with microalbuminuria. *Circulation.* 2004;110:2809-2816.
14. Kettunen J, Demirkan A, Würtz P, Draisma HH, Haller T, Rawal R, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun. 2016;7:11122.
15. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity.* 2005;95:221-227.
16. Kurisu S, Ishibashi K, Kato Y, Mitsuba N, Dohi Y, Nishioka K, et al. Effects of lipid-lowering therapy with strong statin on serum polyunsaturated fatty acid levels in patients with coronary artery disease. *Heart Vessels.* 2013;28:34-38.

17. Jula A, Marniemi J, Rönnemaa T, Virtanen A, Huupponen R. Effects of diet and simvastatin on fatty acid composition in hypercholesterolemic men a randomized controlled trial. *Arterioscler Thromb Vasc Biol.* 2005;25:1952-1959.
18. Benn M, Nordestgaard BG, Jensen GB, Tybjærg-Hansen A. Improving prediction of ischemic cardiovascular disease in the general population using apolipoprotein B The Copenhagen City Heart Study. *Arterioscler Thromb Vasc Biol.* 2007;27:661-670.
19. Kastelein JJ, van der Steeg WA, Holme I, Gaffney M, Cater NB, Barter P, et al. Lipids, apolipoproteins, and their ratios in relation to cardiovascular events with statin treatment. *Circulation.* 2008;117:3002-3009.
20. Kappelle PJ, Dallinga-Thie GM, Dullaart RP, Diabetes Atorvastatin Lipid Intervention (DALI) study group. Atorvastatin treatment lowers fasting remnant-like particle cholesterol and LDL subfraction cholesterol without affecting LDL size in type 2 diabetes mellitus: Relevance for non-HDL cholesterol and apolipoprotein B guideline targets. Biochim Biophys Acta. 2010;1801:89-94.
21. Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. *Lancet.* 2014:384:626-635.
22. Jørgensen AB, Frikke-Schmidt R, West AS, Grande P, Nordestgaard BG, Tybjærg-Hansen A. Genetically elevated non-fasting triglycerides and calculated remnant cholesterol as causal risk factors for myocardial infarction. *Eur Heart J.* 2012;34:1826-1833.
23. Varbo A, Benn M, Tybjærg-Hansen A, Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG. Remnant cholesterol as a causal risk factor for ischemic heart disease. *J Am Coll Cardiol.* 2013;61:427-436.
24. Varbo A, Benn M, Tybjærg-Hansen A, Nordestgaard BG. Elevated remnant cholesterol causes both low-grade inflammation and ischemic heart disease, whereas elevated low-density lipoprotein cholesterol causes ischemic heart disease without inflammation. *Circulation.* 2013;128:1298-1309.
25. Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, et al. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA.* 2009;302:1993–2000.
26. Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J.* 2015;36:539–550.
27. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. *Lancet.* 2012;380:572–580.
28. Keene D, Price C, Shun-Shin MJ, Francis DP. Effect on cardiovascular risk of high density lipoprotein targeted drug treatments niacin, fibrates, and CETP inhibitors: meta-analysis of randomised controlled trials including 117 411 patients. *BMJ.* 2014;349:g4379.
29. Camont L, Chapman MJ, Kontush A. Biological activities of HDL subpopulations and their relevance to cardiovascular disease. *Trends Mol Med.* 2011;17:594-603.
30. Superko HR, Pendyala L, Williams PT, Momary KM, King SB, Garrett BC. High-density lipoprotein subclasses and their relationship to cardiovascular disease. *J Clin Lipidol.* 2012;6:496-523.
31. Mora S, Otvos JD, Rifai N, Rosenson RS, Buring JE, Ridker PM. Lipoprotein particle profiles by nuclear magnetic resonance compared with standard lipids and apolipoproteins in predicting incident cardiovascular disease in women. Circulation. 2009;119:931-939.
32. Musunuru K, Orho-Melander M, Caulfield MP, et al. Ion mobility analysis of lipoprotein subfractions identifies three independent axes of cardiovascular risk. Arterioscler Thromb Vasc Biol. 2009;29:1975-1980.
33. St-Pierre AC, Cantin B, Dagenais GR, Mauriège P, Bernard PM, Després JP, et al. (2005). Low-density lipoprotein subfractions and the long-term risk of ischemic heart disease in men. Arterioscler Thromb Vasc Biol. 2005;*25*:553-559.

34. Ip S, Lichtenstein AH, Chung M, Lau J., Balk EM. Systematic review: association of low-density lipoprotein subfractions with cardiovascular outcomes. Ann Intern Med. 2009:150;474-484.
35. Nordestgaard B, Zilversmit D. Large lipoproteins are excluded from the arterial wall in diabetic cholesterol-fed rabbits. J Lipid Res. 1988;29:1491-1500.
36. Kappelle, PJ., van Tol A, Wolffenbuttel BH, Dullaart RP. Cholesteryl ester transfer protein inhibition in cardiovascular risk management: ongoing trials will end the confusion. Cardiovasc Ther. 2011;29:e89-99.
37. Ahnadi CE, Berthezène F, Ponsin G. Simvastatin-induced decrease in the transfer of cholesterol esters from high density lipoproteins to very low and low density lipoproteins in normolipidemic subjects. Atherosclerosis. 1993;99:219-228.
38. de Vries R, Dikkeschei BD, Sluiter WJ, Dallinga-Thie GM, van Tol A, Dullaart, RP. Statin and fibrate combination does not additionally lower plasma cholesteryl ester transfer in type 2 diabetes mellitus. Clin Lab., 2012;58:1231-1239
39. de Vries R, Kerstens MN, Sluiter,WJ, Groen AK, van Tol A, Dullaart RPF. Cellular cholesterol efflux to plasma from moderately hypercholesterolaemic type 1 diabetic patients is enhanced, and is unaffected by simvastatin treatment. Diabetologia. 2005;48:1105-1113.
40. Packard CJ, Shepherd J (1997). Lipoprotein heterogeneity and apolipoprotein B metabolism. Arterioscler Thromb Vasc Biol. 1997;17: 3542-3556.
41. Reimann FM, Winkelmann F, Fellermann K, Stange EF. Reduced cholesterol esterification in CaCo-2 cells by indirect action of pravastatin. *Atherosclerosis.* 1996;125:63-70.
42. Jasińska M, Owczarek J, Orszulak-Michalak D. Statins: a new insight into their mechanisms of action and consequent pleiotropic effects. Pharmacol Rep. 2007;59:483-499.
43. Nozue T, Michishita I. Statin treatment alters serum n-3 to n-6 polyunsaturated fatty acids ratio in patients with dyslipidemia. Lipids Health Dis. 2005;14:67.
44. Nozue T, Yamamoto S, Tohyama S, Fukui K, Umezawa S, Onishi Y, et al. Comparison of effects of serum n-3 to n-6 polyunsaturated fatty acid ratios on coronary atherosclerosis in patients treated with pitavastatin or pravastatin undergoing percutaneous coronary intervention. Am J Cardiol. 2013:111;1570-1575.
45. Kassai A, Illyés L, Mirdamadi HZ, Seres I, Kalmár T, Audikovszky M, et al. The effect of atorvastatin therapy on lecithin: cholesterol acyltransferase, cholesteryl ester transfer protein and the antioxidant paraoxonase. *Clin Biochem. 2007;40*:1-5.
46. Mayes PA. Intermediary metabolism of fructose. Am J Clin Nutr. 1993; 58:754S-765S
47. Schaefer EJ, McNamara JR, Tayler T, Daly JA, Gleason JL, Seman L, et al. Comparisons of effects of statins (atorvastatin, fluvastatin, lovastatin, pravastatin, and simvastatin) on fasting and postprandial lipoproteins in patients with coronary heart disease versus control subjects. Am J Cardiol. 2004:93;31-39.

Chapter 8

Summary and Discussion

PART I GENOMICS IN CVD

In the first part of this thesis (**Chapter 2**) we explored the genomics of resting heart rate and mortality. Previous studies showed an association between heart rate and life expectancy or risk, but do not provide sufficient evidence for a (shared) causal relationship. We found evidence that genetic variants associated with higher resting heart rate also confer a risk for death. Additionally, we found that these variants are associated with potential measured (body mass index, systolic and diastolic blood pressure, etc.) and unmeasured confounders. We aimed to adjust for this, by allowing genetic variants to have pleiotropic effects, removing genetic variants associated with other traits, or using estimates derived from healthy participants and 130,795 independent participants. All of our analyses consistently suggest that heart rate is linked to mortality, and therefore life-expectancy. There are two different possibilities: either the genetic variants exert their effect on mortality directly via heart rate as a mediator or, alternatively, the genetic variants share underlying biology, resulting in both increased heart rate and increased mortality risk. Basic cellular biology underlying heart rate may be involved, as may vulnerability to cardiac arrhythmias causing (sudden) death, which could contribute to all classifications of death and eventually prove relevant to a plethora of non-cardiac diseases and conditions. The fact that predominantly cardiac candidate genes were identified at the identified loci and the colocalization of DNase hypersensitivity sites in cardiac tissue further supports this theory. However, alternative mechanisms involving basic metabolic rate, energetics, and free radicals could result in cumulative general damage and affect lifespan. In addition to an interpretation of causation, a number of candidate genes have a known function relevant for cardiac conditions. However, we have not proven that this mechanism explains the association with heart rate for any of these genes. Further experimental validation of each locus is needed to identify the underlying biological mechanisms. Additionally, we did not identify differences in a specific cause of death (e.g. cardiovascular system), which might be expected. Associations between resting heart rate and specific causes of death should be investigated further as more long-term follow-up data becomes available. In **Chapter 3**, we studied genetically determined telomere length (TL) and its link with cardiovascular disease (CVD) and cancer. Phenotypic associations between short TL and various CVD (e.g. atherosclerosis, myocardial infarction, and heart failure) have been reported in previous studies[1,2]. We provide evidence for a causal link between genetically determined TL and the development of hypertension, CVD and cancer using a Mendelian randomization approach. However, the exact molecular mechanisms underlying these associations remain to be elucidated. For example, in hypertension it may be that TL in vascular endothelial cells is associated with ageing processes, such as thickening of the arterial wall, increased deposition of collagen and loss of elastin. These changes lead to increased arterial

stiffness and consequent hypertension. The results on cancer also suggest causal links with TL. Further studies are required to analyze the separate CVD conditions, such as specific forms of atherosclerosis or heart failure and specific cancer types. In **Chapter 4**, we explore the genomics of coronary artery disease (CAD). In order to improve our biological understanding of CAD and facilitate the identification of therapeutic targets, and gain an insight into the causal relationships between other cardiovascular phenotypes, we used a two stage approach, adding new cases and controls to the data from the CARDIoGRAMplusC4D consortium. We identified 15 novel genome-wide significant loci, adding a substantial number to the 57 previously known loci[3]. We observed that the genetic risk score of CAD predicts a range of cardiovascular phenotypes (e.g. heart failure and atrial fibrillation), consistent with clinical practice. Further functional experiments are warranted in order to establish further evidence for the true causal genes and mechanisms underlying each association.

PART II METABOLOMICS IN CVD

In the second part of this thesis we investigated the effect of metabolites on left ventricular ejection fraction and infarct size, as well as the effect of metformin treatment on metabolites in patients after a myocardial infarction using data from the GIPS-III study[4]. In **Chapter 5**, we investigated lipoprotein subfractions in the GIPS-III study by applying nuclear magnetic resonance (NMR) spectroscopy using the LipoScience platform. In subjects with impaired glucose tolerance, metformin administration modestly reduces the low density lipoprotein (LDL) particle concentration and concomitantly decreases small dense LDL particles and increases small and large high density lipoprotein (HDL) particles[5]. Our study showed that metformin treatment initiated directly after the acute phase of MI modestly decreases LDL cholesterol and LDL size. It is known that in CAD patients, HDL cholesterol and smaller-sized HDL may confer higher left ventricular ejection fraction (LVEF)[6,7]. In the GIPS-III study, metformin did not affect LVEF[4] and we therefore considered the metformin- and placebo-receiving participants together. We show that elevated medium very low density lipoprotein and small HDL particle concentrations 24 hours post-MI may confer beneficial associations with increased LVEF and decreased infarct size. In **Chapter 6**, we studied the same cohort using a different NMR spectroscopy platform (Computational Medicine), which includes amino acids. In addition, we tried to identify subgroups of patients in whom metformin was effective using metabolic profiling. We found that after 4 months of metformin treatment, alanine levels and phospholipids/total lipids ratio in very large HDL differed significantly. Furthermore, we found that higher triglyceride levels in HDL measured 24 hours post-MI were associated with favorable outcome as inferred from higher LVEF and smaller infarct size 4 months

post-MI. We could not identify metabolic profiles associated with treatment benefits of metformin. Data from ongoing follow-up will be available in the near future, which will allow us to study the effects of metformin treatment and drug metabolite interactions on long-term outcome. We should take into account some aspects of the GIPS-III study design to place the results of both Chapter 5 and 6 in perspective[8]. Due to the study design, subjects were included shortly after arrival at the hospital, plasma lipid measurements were not carried out in the fasting state and, for logistic reasons, non-fasting samples were also obtained during follow-up. It should be stressed that all patients received intravenous heparin before undergoing percutaneous coronary interventions. Heparin stimulates lipolysis[9] and hence acutely reduces plasma triglyceride levels[10]. Association of metabolites with hard clinical end-points could not be assessed because only 2% of the participants experienced recurrent major adverse cardiac events, and none of the participants died during the 4 months follow-up[4]. In **Chapter 7**, we studied the effect of statin treatment on metabolomics by using data from the PREVEND-IT study. It is well-known that statins reduce low-density lipoprotein cholesterol (LDL-C)[11] but also alter a wide range of lipids and fatty acids (FA)-related metabolites[12,13]. We found that pravastatin treatment for 3 months at a dose of 40 mg once daily lowered LDL-C, remnant cholesterol and apolipoprotein B with similar effect magnitudes. Furthermore, we observed globally lowered levels of lipid subfractions, with the exception of HDLs subfractions, which displayed a more complex response pattern. Metabolic profiling is emerging as a powerful tool to dissect a drug's efficacy profile, providing important information that may be used to improve current treatments.

FUTURE PERSPECTIVES

Goals of the cardiovascular genomics field are to understand biological mechanisms and use this knowledge to personalize medicine. Gathering knowledge of molecular pathways can lead to improved therapeutics, which can be targeted based on individual genotype independently of the phenotype. Numerous CVD loci have been discovered in recent years, to which this thesis has contributed for CAD and resting heart rate phenotypes. The number of loci will expand rapidly in the future as genome-wide DNA sequencing becomes more cost-efficient. This may allow us to detect rare frequency variants, for example, which we did not find for CAD and resting heart rate. In the future, functional studies will be required to map molecular and cellular pathways with greater accuracy, which will provide opportunities for the development of appropriately targeted therapies for CVD. Performing functional follow-up of candidate genes using technologies as CRIPSPR in model organisms or relevant human cell lines may provide us with a way to further unravel the mechanisms of how resting heart rate is regulated.

In the future, further studies are needed to study the relationship of resting heart rate with mortality in greater depth by identifying the main driver of this association. This could be done by using new follow-up data from the UK Biobank and independent cohorts, which may be better powered to determine specific causes of death driving the association with mortality. Furthermore, this may also help us establish additional relationships between genetically determined TL and more specific CVDs.

In parallel to the evolution of genomics, metabolomics is also rapidly emerging as a powerful technological platform for understanding mechanisms underlying common chronic diseases such as CVD. Metabolomics measure the product upstream genetic, transcriptomic, and proteomic variation. In the future, metabolomics and genomics should be integrated more closely in order to generate hypotheses by using metabolites underlying a specific genetic variant or gene. The integration of different 'omics' (genomics, transcriptomics, metabolomics, proteomics, etc.) should be pursued in order to more accurately characterise causal relationships, contribute to less biased hypothesis-generating studies to identify pathways of disease, and improve risk prediction by identifying disease markers and potential new targets for medication.

REFERENCES

1. Wong LS, de Boer RA, Samani NJ, van Veldhuisen DJ, van der Harst P. Telomere biology in heart failure. *Eur J Heart Fail*. 2008;10(11):1049-1056.
2. Samani NJ, van der Harst P. Biological ageing and cardiovascular disease. *Heart*. 2008;94(5):537-539.
3. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121-1130.
4. Lexis CP, van der Horst IC, Lipsic E, et al. Effect of metformin on left ventricular function after acute myocardial infarction in patients without diabetes: The GIPS-III randomized clinical trial. *JAMA*. 2014;311(15):1526-1535.
5. Goldberg R, Temprosa M, Otvos J, et al. Lifestyle and metformin treatment favorably influence lipoprotein subfraction distribution in the diabetes prevention program. *J Clin Endocrinol Metab*. 2013;98(10):3989-3998.
6. Kempen HJ, van Gent CM, Buytenhek R, Buis B. Association of cholesterol concentrations in low-density lipoprotein, high-density lipoprotein, and high-density lipoprotein subfractions, and of apolipoproteins AI and AII, with coronary stenosis and left ventricular function. *J Lab Clin Med*. 1987;109(1):19-26.
7. Yang N, Feng JP, Chen G, et al. Variability in lipid profile among patients presented with acute myocardial infarction, unstable angina and stable angina pectoris. *Eur Rev Med Pharmacol Sci*. 2014;18(24):3761-3766.
8. Lexis CP, van der Horst IC, Lipsic E, et al. Metformin in non-diabetic patients presenting with ST elevation myocardial infarction: Rationale and design of the glycometabolic intervention as adjunct to primary percutaneous intervention in ST elevation myocardial infarction (GIPS)-III trial. *Cardiovasc Drugs Ther*. 2012;26(5):417-426.
9. Riemens SC, van Tol A, Scheek LM, Dullaart RP. Plasma cholesteryl ester transfer and hepatic lipase activity are related to high-density lipoprotein cholesterol in association with insulin resistance in type 2 diabetic and non-diabetic subjects. *Scand J Clin Lab Invest*. 2001;61(1):1-9.
10. Brunner MP, Shah SH, Craig DM, et al. Effect of heparin administration on metabolomic profiles in samples obtained during cardiac catheterization. *Circ Cardiovasc Genet*. 2011;4(6):695-700.
11. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: A meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet*. 2010;376(9753):1670-1681.
12. Wurtz P, Wang Q, Soininen P, et al. Metabolomic profiling of statin use and genetic inhibition of HMG-CoA reductase. *J Am Coll Cardiol*. 2016;67(10):1200-1210.
13. Kurisu S, Ishibashi K, Kato Y, et al. Effects of lipid-lowering therapy with strong statin on serum polyunsaturated fatty acid levels in patients with coronary artery disease. *Heart Vessels*. 2013;28(1):34-38.

Chapter 9

Appendix

NEDERLANDSE SAMENVATTING

Inleiding

Hart- en vaatziekten vormen samen de belangrijkste doodsoorzaak wereldwijd. Veelvoorkomende hart- en vaatziekten zijn coronaire hartziekte, hypertensie, cerebrovasculaire ziekten en perifeer vaatlijden. In de afgelopen 50 jaar is er aanzienlijke vooruitgang geboekt in de definitie en identificatie van de risicofactoren en behandeling van hart- en vaatziekten, waaronder de ontwikkeling van van passende medische en interventiebehandelingen, zoals percutane coronaire interventies en het gebruik van β-blokkers. Al deze maatregelen hebben geleid tot een daling van de cardiovasculaire sterfte. Ondanks deze inspanningen zijn de mechanismen die aan de pathofysiologie van hart- en vaatziekten ten grondslag liggen, nog steeds gedeeltelijk begrepen.

Het verkrijgen van nieuwe inzichten in biologische en causale verbanden vergroot ons begrip van de pathofysiologie van hart- en vaatziekten en maakt de ontwikkeling van nieuwe therapeutische strategieën en risicostratificatie mogelijk. In dit proefschrift werden genomics (**Deel I**) en metabolomics (**Deel II**) toegepast om biologisch en klinisch inzicht te krijgen in de eigenschappen van hart- en vaatziekten. Hiervoor werden drie overkoepelende methodologieën gebruikt: a) Genoombrede associatiestudies ('genome-wide association studies', GWAS) om nieuwe genetische regio's en genen te identificeren om de pathofysiologie van hart- en vaatziekten beter te begrijpen; b) Mendeliaanse randomisatie om potentiële oorzaken van een ziekte te identificeren; en c) Metabolomics om de producten van gentranscriptie te bestuderen en nieuwe biomarkers te identificeren.

Deel I Genomics van hart- en vaatziekten

Het eerste ontwerp van het menselijke genoom (Human Genome Project) werd vijftien jaar geleden geproduceerd, wat leidde tot een beter begrip van de genetische bijdrage van gemeenschappelijke varianten aan hart- en vaatziekten. Bij het Human Genome Project waren genen geassocieerd met hart- en vaatziekten via Mendeliaanse associatie (ook wel monogentische overervering), welke relatief zeldzaam zijn en slechts een klein deel van de klinische hart- en vaatziekten vormen. Voorbeelden zijn: familiaire hypercholesterolemie, dilaterende en hypertrofische cardiomyopathie, lange-QT syndroom en aorta aneurysma's. Het merendeel van de hart- en vaatziekten is echter polygeen, met veel erfelijke en omgevingsfactoren. Voorafgaand aan de voltooiing van het ontwerp van het menselijk genoom waren inspanningen om de genetische oorzaken van polygene hart- en vaatziekten te identificeren grotendeels mislukt. Door genoom-brede associatiestudies, die genetische varianten van het genoom testen voor hun associatie met een ziekte of eigenschap, werden honderden loci voor talrijke hart- en vaatziekten en eigenschappen ontdekt/geïdentificeerd.

Het doel van dit proefschrift is het uitbreiden van dit onderzoeksgebied en het creëren van nieuwe inzichten in de genomics van het cardiovasculaire systeem. We gebruiken GWAS om nieuwe genetische varianten te identificeren en daardoor onze kennis verder te ontwikkelen. Verschillende bioinformatische methoden werden toegepast op de gevonden genetische varianten om nieuwe biologische mechanismen te vinden die een rol spelen in hart- en vaat ziekten. Tevens passen we Mendeliaanse randomisatie analyses toe, waarbij gebruik wordt gemaakt van de genetische varianten die zijn geidentificeerd door GWAS om het causale effect van risicofactoren op ziekteontwikkeling en sterfte te schatten.

In **hoofdstuk 2** verdiepten we ons in de genetica van hartslag in rust en de relatie met mortaliteit. Eerdere studies hebben een verband laten zien tussen hartslag en levensverwachting, maar bieden onvoldoende bewijs voor een (gedeelde) causale relatie. In dit hoofdstuk hebben we bewijzen gevonden dat genetische varianten, die verband houden met een hogere rusthartslag, zorgen voor een hoger risico op overlijden. Verder vonden we dat onze geïdentificeerde varianten geassocieerd zijn met potentiele gemeten (Body Mass Index, systolische en diastolische bloeddruk, enz.) en ongemeten parameters. In deze studie hebben we ernaar gestreefd om hiervoor te corrigeren. De resultaten wijzen erop dat de rusthartslag is gekoppeld aan de levensverwachting van de mens. In **hoofdstuk 3** bestudeerden we de associatie van genetisch bepaalde telomeerlengte met cardiovasculaire aandoeningen en kanker. Fenotypische associaties tussen korte telomeerlengte en verschillende cardiovasculaire aandoeningen (bijvoorbeeld atherosclerose, een hartinfarct en hartfalen) zijn gerapporteerd in eerdere studies. We leveren bewijs voor een oorzakelijk verband tussen genetisch bepaalde telomeerlengte en de ontwikkeling van hypertensie, hart- en vaatziekten en kanker met behulp van een Mendeliaanse randomisatiebenadering. In **hoofdstuk 4** onderzochten we de genetica van coronaire hartziekte. Om ons biologisch begrip van coronaire hartziekte te vergroten, identificatie van therapeutische doelen te vergemakkelijken en inzicht te krijgen in de causale relaties tussen andere cardiovasculaire fenotypes, hebben we nieuwe gevallen en controles toegevoegd aan de gegevens van een bestaand consortium (CARDIoGRAMplusC4D). We identificeerden vijftien nieuwe genoomwijde significante loci, die een aanzienlijk aantal bij de 57 voorheen bekende loci toevoegden. Tevens constateerden we dat de genetische risicoscore van coronaire hartziekte een reeks van andere cardiovasculaire fenotypen zoals hartfalen en atriumfibrilleren voorspelt, consistent met de klinische praktijk.

Deel II Metabolomics van hart- en vaatziekten

Terwijl genetische varianten essentiële onderdelen van erfelijkheid vormen, zijn ze relatief 'statische' componenten. Voor chronische ziekten, zoals hart- en vaatziekten,

kan een nauwkeuriger onderzoek van het ziekteproces waardevol zijn om biomarkers te identificeren en mechanismen en pathofysiologie van hart- en vaatziekten begrijpen. Metabolomics, het bestuderen van de producten van gentranscriptie, biedt nieuwe mogelijkheden om nog onbekende biomarkers te vinden. Het biedt ook nieuwe mogelijkheden om de biologische mechanismen van ziekte te bestuderen, omdat metabolieten de producten van gentranscriptie vertegenwoordigen. Metabolomics is een relatief nieuw vakgebied in 'omics' wetenschappen, die gebruik maakt van high-throughput technologieën, zoals nucleaire magnetische resonantie (NMR) spectroscopie, om tegelijkertijd een groot aantal kleine moleculen in verschillende weefsels te kwantificeren. Metabolische profilering is succesvol geweest bij het verbeteren van de diagnose en voorspelling van cardiovasculaire events en bij het differentiëren van hartfalenpatiënten met gezonde controles. Metabolische profilering kan ons dus potentieel helpen bij het identificeren van nieuwe biomarkers.

In deel twee van dit proefschrift bestuderen we de veranderingen die betrokken zijn bij metabolieten bij patiënten met een acuut myocardinfarct en de effecten van statinetherapie op metaboliet profielen met behulp van NMR spectroscopie. In **hoofdstuk 5** onderzochten we lipoproteïne subfracties in de GIPS-III studie door toepassing van NMR spectroscopie met behulp van het LipoScience platform. Uit onze studie bleek dat behandeling met metformine direct na de acute fase van een hartinfarct het lage dichtheids lipoproteïne (LDL)-cholesterol en de LDL-grootte verminderde. Het is bekend dat bij patiënten met coronaire hartziekte het hoge dichtheids lipoproteïne (HDL)-cholesterol en de kleinere HDL deeltjes geassocieerd zijn met een hogere linker ventrikel ejectiefractie. In de GIPS-III studie had metformine geen invloed op linker ventrikel ejectiefractie en daarom hebben we de deelnemers met metformine en placebo samengevoegd in onze analyse. We tonen aan dat verhoogde concentraties van kleinere HDL-deeltjes 24 uur na een acuut myocardinfarct een hogere hogere linker ventrikel ejectiefractie voorspellen. Verder laten we zien dat verhoogde concentratrie van medium zeer lage dichtheidslipoproteïne deeltjes een gunstige associatie hebben met een kleinere infarctgrootte. In **hoofdstuk 6** bestudeerden we hetzelfde cohort met behulp van een ander NMR spectroscopieplatform (Computational Medicine), waarin ook aminozuren werden gemeten. We vonden dat na vier maanden metforminebehandeling, alanine niveaus en de verhouding van fosfolipiden tot totale lipiden in zeer grote HDL deeltjes significant verschilden. Verder vonden we dat de HDL triglyceride concentratie na een infarct de linker ventrikel ejectiefractie en infarctgrootte voorspelde. Met behulp van metabolische profilering konden we geen subgroepen van patiënten te identificeren bij wie metformine een positief effect had op de linker ventrikel ejectiefractie. In **hoofdstuk 7** hebben we het effect van statinebehandeling op het metabolische profiel bestudeerd door gebruik te maken van gegevens uit de PREVEND-IT studie. Het is algemeen bekend dat statines LDL-cholesterol verminderen, maar ook een groot aantal

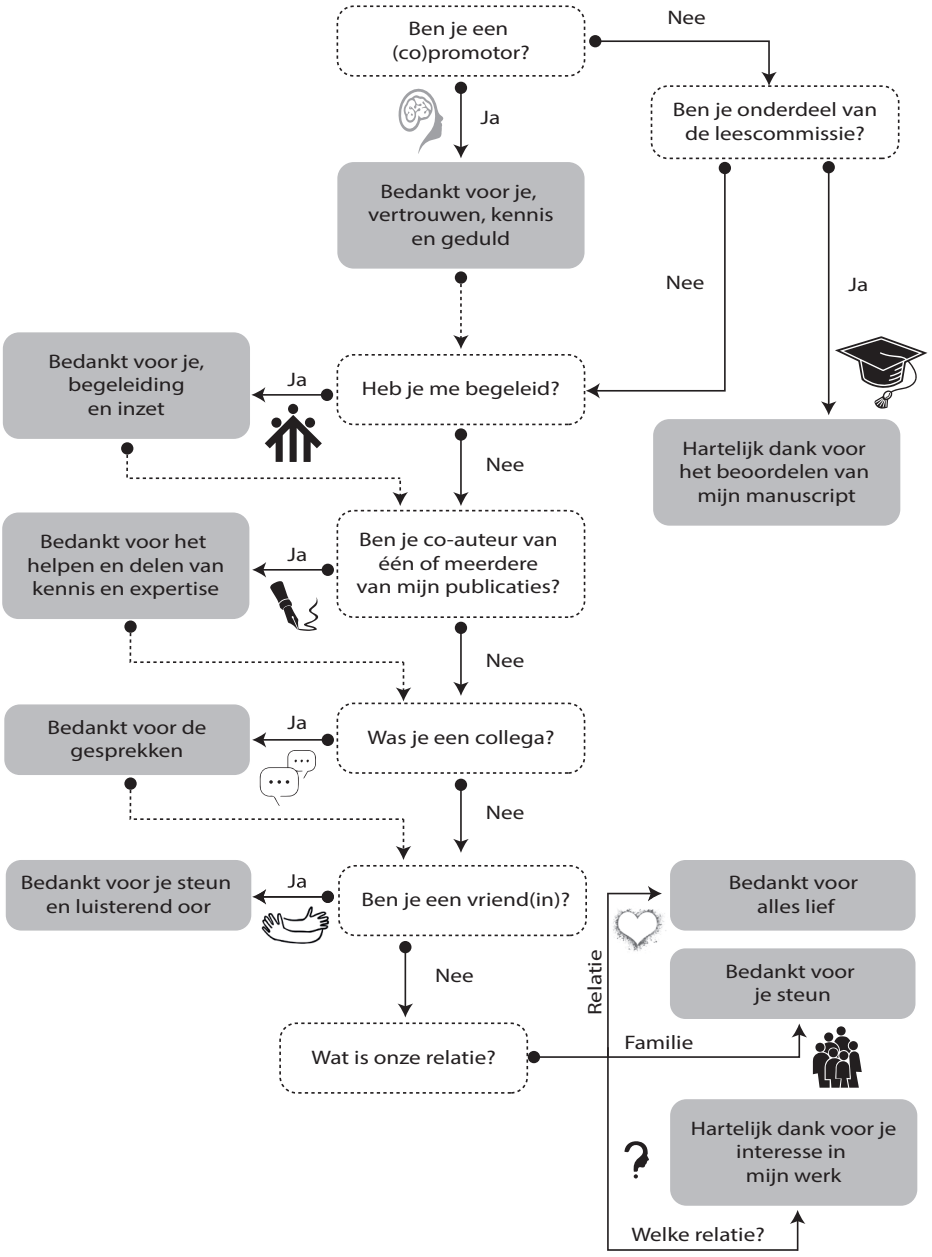
lipiden en vetzuren gerelateerde metabolieten veranderen. We vonden dat behandeling met pravastatine, gedurende drie maanden, LDL cholesterol, restant cholesterol en apolipoproteïne B met vergelijkbare effectgroottes verlaagden. Verder hebben we globaal verlaagde niveaus van lipide subfracties waargenomen, met uitzondering van HDL subfracties, die een complexer responspatroon vertoonden.

CURRICULUM VITAE

Ruben Nathaniël Eppinga was born on April 15th, 1984 in Drachten, the Netherlands, as the first-born child of Hille and Sjoukje. He attended primary school in Beetsterzwaag at De Paedwizer. After primary school he started secondary school (MAVO, the Rehoboth) in Drachten. After finishing his secondary school Ruben started Technical Informatics at the ROC Friesepoort in Drachten. From this time, he started studying Bioinformatics at the Hanze Hogeschool in Groningen and obtained his propaedeutics. He travelled through Australia with friends for approximately half a year. When returned home he studied for entrance examinations for medical school. In 2006 Ruben started to study medicine at the University of Groningen (RUG). In 2009 he finished his Junior Scientific Masterclass Bachelor 'Honours'-programme and obtained his 'Honours'-degree. In 2013 he graduated and obtained his MD degree (cum laude) at the University of Groningen. Subsequently, he worked for half a year as a resident Intensive Care at the Medical Centre Leeuwarden. Ruben started his PhD project in 2013 at the University Medical Center Groningen, under the supervision of cardiologist prof. dr. Pim van der Harst. He could combine his background in Bioinformatics and personal interest in big data in his research project which resulted in this thesis. Ruben defends his dissertation entitled 'Genomics and Metabolomics insights into Cardiovascular Disease' on May 16th 2018 in Groningen at the RUG.

PUBLICATIONS

1. Said MA*, Eppinga RN*, Lipsic E, et al. Relationship of Arterial Stiffness Index and Pulse Pressure with Cardiovascular Disease and Mortality. ***J Am Heart Assoc***. 2018 Jan 22;7(2).
2. Kofink D*, Eppinga RN*, van Gilst WH, et al. Statin effects on metabolic profiles: data from the PREVEND IT trial. ***Circ Cardiovasc Genet.*** 2017 Dec;10(6).
3. Said MA*, Eppinga RN*, Hagemeijer Y, et al. Telomere length and risk of cardiovascular disease and cancer: a Mendelian Randomization in 134,773 individuals. ***J Am Coll Cardiol.*** 2017 Jul 25;70(4):506-507.
4. Eppinga RN*, Hartman MH*, Verweij N, et al. Refining thromboembolic risk in the general population. ***J Am Coll Cardiol.*** 2017 Jul 18;70(3):403-404.
5. Verweij N, Eppinga RN, Hagemeijer Y, van der Harst P. Identification of 15 novel risk loci for coronary artery disease and genetic risk of recurrent events, atrial fibrillation and heart failure. ***Sci Rep.*** 2017 Jun 5;7(1):2761.
6. van den Berg ME, Warren HR, Cabrera CP, …, Eppinga RN, et al. Discovery of novel heart rate-associated loci using the Exome Chip. ***Hum Mol Genet.*** 2017 Apr 3.
7. Eppinga RN*, Kofink D*, Dullaart RP, …, van der Harst P. Effect of Metformin on Metabolites and Relation With Myocardial Infarct Size and Left Ventricular Ejection Fraction After Myocardial Infarction. ***Circ Cardiovasc Genet.*** 2017 Feb;10(1).
8. Schmidt AF, Swerdlow DI, Holmes MV, ..., Eppinga RN, et al. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. ***Lancet Diabetes Endocrinol.*** 2017 Feb;5(2):97-105
9. Eppinga RN, Hagemeijer Y, Burgess S, Hinds DA, Stefansson K, Gudbjartsson DF, van Veldhuisen DJ, Munroe PB, Verweij N, van der Harst P. Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. ***Nat Genet.*** 2016 Dec;48(12):1557-1563.
10. Hartman MH, Eppinga RN, Vlaar PJ, …, van der Harst P. The contemporary value of peak creatine kinase-MB after ST-segment elevation myocardial infarction above other clinical and angiographic characteristics in predicting infarct size, left ventricular ejection fraction, and mortality. ***Clin Cardiol.*** 2016 Dec 27.
11. Zhang W, Jernerén F, Lehne BC, ..., Eppinga RN, et al. Genome-wide association reveals that common genetic variation in the kallikrein-kinin system is associated with serum L-arginine levels. ***Thromb Haemost.*** 2016 Nov 30;116(6):1041-1049.
12. Eppinga RN, Hartman MH, van Veldhuisen DJ, et al. Effect of Metformin Treatment on Lipoprotein Subfractions in Non-Diabetic Patients with Acute Myocardial Infarction: A Glycometabolic Intervention as Adjunct to Primary Coronary Intervention in ST Elevation Myocardial Infarction (GIPS-III) Trial. ***PLoS One.*** 2016 Jan 25;11(1):e0145719.
13. Lanting CP, de Kleine E, Eppinga RN, et al. Neural correlates of human somatosensory integration in tinnitus. ***Hear Res.*** 2010 Aug;267(1-2):78-88.


Ben je een (co)promotor?
Nee
Ja
Ben je onderdeel van de leescommissie?
Bedankt voor je, vertrouwen, kennis en geduld
Nee
Ja
Bedankt voor je, begeleiding en inzet
Ja
Heb je me begeleid?
Hartelijk dank voor het beoordelen van mijn manuscript
Nee
Bedankt voor het helpen en delen van kennis en expertise
Ja
Ben je co-auteur van één of meerdere van mijn publicaties?
Nee
Bedankt voor de gesprekken
Ja
Was je een collega?
Nee
Bedankt voor je steun en luisterend oor
Ja
Ben je een vriend(in)?
Nee
Relatie
Bedankt voor alles lief
Bedankt voor je steun
Wat is onze relatie?
Familie
Hartelijk dank voor je interesse in mijn werk
Welke relatie?

DANKWOORD

Inleiding

Het dankwoord is een veel gelezen en goed bestudeerd onderdeel van een proefschrift. Naarmate de promotie langer duurt, wordt de kans groter dat er mensen worden vergeten, wat in de praktijk voor vervelende situaties kan zorgen. Om de kans hiertoe aanzienlijk te verkleinen, wordt er in dit proefschrift gebruik gemaakt van een interactief dankwoord.

Methode en resultaten

Het interactieve dankwoord is afgebeeld in een flow-diagram (Figuur 1). De lezer dient te beginnen met de vraag "Ben je een (co)promotor?" en dient via het volgen van de pijlen uit te komen bij de volgende vraag (witte box) of een antwoord (grijze box). Het kan voorkomen dat de lezer in meerdere grijze boxen terecht komt en meerdere woorden van dank zal ontvangen. Gezien het feit dat een vergelijkbaar flow-diagram tot op heden nog maar een keer is afgebeeld[1] zijn er nog niet genoeg validatiestudies gedaan waardoor er extra tekst noodzakelijk is naast het flow-diagram.

Als allereerste wil ik mijn (co)promotoren Pim, Wiek en Niek bedanken. Pim, hartelijk dank voor je vertrouwen, tijd en energie die je in mijn promotie hebt gestoken. Je doorzettingsvermogen en inzet tijdens je eigen vakantie (in Friesland op de camping met the kids) voor het heart rate stuk kan ik mij nog goed herinneren en waardeer ik enorm. Wiek, bedankt voor je kritische blik en vragen tijdens o.a. de presentaties bij de experimentele cardiologie. Je was altijd rustig en helder, maar niet minder scherp in je observaties. Niek, bedankt voor je tijd, geduld, het delen van je kennis en heldere uitleg. Je bent een meester in "even een stukje code bouwen"; jij weet dit in een mum van tijd klaar te spelen, terwijl menig ander hier tijden mee bezig is. Je hebt je eigen stelling: "Use a computer as a saw; let it do the work" altijd hoog weten te houden.

Beste leden van de beoordelingscommissie, prof. dr. P.I.W. de Bakker, prof. dr. B.H.Ch. Stricker en prof. dr. M.P. van den Berg hartelijk dank voor het beoordelen van van mijn proefschrift.

Beste collega's en coauteurs, hartelijk dank voor de inhoudelijke discussies en voor het helpen tot stand komen van dit proefschrift. Zonder jullie was dit proefschrift nooit geworden wat het nu is. Een speciale dank voor mijn oud-collega Yanick, die samen met mij tijdens de analyses midden in de nacht met behulp van alle computers de kamer tot meer dan dertig graden op wist te stoken.

Paranimfen, Minke en zusje Hanna, bedankt voor jullie inzet en tijd. Het is een fijne gedachte dat, als ik het niet red, jullie mijn verdediging willen overnemen.

Lieve Hille en Sjoukje, familie en vrienden, dank voor jullie interesse, steun en luisterend oor. Ook als ik er even geen zin meer in had, wisten jullie mij toch te motiveren om door te gaan en met name ook de nodige afleiding te creëren.

Lieve Rixte, bedankt voor je onvoorwaardelijke liefde en steun, ik heb zin in de toekomst samen.

Conclusie

Zonder de collega's, coauteurs, vrienden en familie was dit proefschrift nooit tot stand gekomen. Iedereen die op welke manier dan ook heeft bijgedragen aan mijn proefschrift wil ik via deze weg bedanken.

LITERATUUR

1. Kockelkorn, T.T.J.P. (2017). Interactive texture analysis in chest CT scans (Doctoral dissertation of Master's thesis). Retrieved from https://dspace.library.uu.nl/bitstream/handle/1874/343111/Kockelkorn.pdf?sequence=1&isAllowed=y